Date: May 7, 1999

From: Ellis F. Unger, M.D., DCTDA, OTRR

Subject: Clinical Review of sBLA 98-0737
Chiron Corp.; Interferon β-1b (Betaseron®)

Through: Marc Walton, M.D., Ph.D., Chief, Medicine Branch, DCTDA
Karen Weiss, M.D., Director, DCTDA

To: sBLA 98- File

This document is the Medical Officer Clinical Review for sBLA-98-0737

Sponsor: Chiron Corp.
Product: Interferon β-1b (Betaseron®)
Proposed Indication: treatment of secondary progressive forms of Multiple Sclerosis
Proposed Regimen: 0.25 mg SC QOD, indefinitely
Date of Submission:
Original: June 29, 1998

Supplemental Information:
October 7, 1998
October 19, 1998
December 2, 1998
January 12, 1999
January 19, 1999
January 22, 1999
February 17, 1999

Table of Contents

Overview

Interferon β-1b (Betaseron®) was approved by the FDA for treatment of patients with the relapsing-remitting form of multiple sclerosis (RRMS) in 1993. Chiron, Inc. has submitted a supplemental BLA, and is seeking approval to expand the indication section to include the treatment of secondary progressive multiple sclerosis (SPMS). The sponsor was granted Fast Track status on May 1, 1998.

The proposed claims are:

- slowing the rate of disease progression
- decrease in relapse rate
- decrease in relapse severity
- reducing lesion load, as assessed by MRI, thought to reflect underlying disease activity

Scope of this review

The focus of this document is a single study, ME 93079, a double-blind placebo-controlled multicenter study to evaluate the safety and efficacy of interferon β-1b administered subcutaneously to outpatients with secondary progressive multiple sclerosis. The study was conducted between September 1994 and March 1998 by ----------, with data analysis and interpretation by ----------. Licensure of Betaseron for relapsing remitting multiple sclerosis (PLA 92-0495) was based on study TB01-35686/35886, which is not reviewed herein.

[

] studies can not directly address the safety or efficacy of the product, and will not be addressed extensively in this document.

Funding: ----------

Abbreviations used in this review

CMH	Cochran-Mantel-Haenszel
CNS	central nervous system
ECG	electrocardiogram
EDSS	Expanded Disability Status Scale
EU	European Union
FS	Functional Systems
GEMS	Global Evaluation of Multiple Sclerosis
ITT	intent-to-treat
IFN	interferon
IFNs	interferons
MADRS	Montgomery and Asberg Depression Rating Scale
MRI	magnetic resonance imaging
mIU	million International Units
MS	Multiple sclerosis

NAB	neutralizing antibodies
NSAID	non-steroidal anti-inflammatory drugs
QID	four times daily
QOL	quality of life
RRMS	Relapsing-Remitting Multiple sclerosis
SIP	Sickness Impact Profile
SPMS	Secondary Progressive Multiple sclerosis
TID	three times daily
TIW	three times each week

Introduction

Multiple Sclerosis

Multiple sclerosis (MS) is a debilitating remitting and relapsing autoimmune disease characterized by inflammation and demyelination of nerve tracts within the white matter of the central nervous system (CNS). The disease affects 300,000 patients in the US, with an annual incidence of approximately 9000. It is a disease of young adults, with a median age of onset of 28 years. For reasons that remain unknown, MS is more common in the higher latitudes of both the northern and southern hemispheres.

The etiology of MS is unknown; however, it is widely considered to be an autoimmune disease, where the ongoing destruction of CNS white matter is due to immune system attack directed against CNS myelin. Myelin basic protein (MBP), found in CNS myelin, is a putative agent in sustaining this autoimmune process. Most proposed therapies for MS have focused on modulating immune system function.

Pathologically, MS is characterized by multiple lesions throughout the CNS, with a predilection for the optic nerves, periventricular portions of the lateral ventricles, brain stem, cerebellum and dorsal aspect of the spinal cord. Initially, the inflammatory process involves the myelin sheaths, with preservation of the axons. Partial or complete resolution is due to a decrease in inflammation and myelin repair. With time, the disorder becomes more widely distributed throughout the central nervous system, and repair is less complete. The lesions become more destructive with gliosis and axonal degeneration.

The CNS lesions are manifested clinically as focal deficits or exacerbations. Most common are paresthesias, diplopia, impaired vision, sensory loss, motor weakness, tremor, ataxia, bladder and bowel dysfunction and neuropsychiatric disorders. In the early stages, ≈85% of patients experience a relapsing-remitting course, characterized by episodic and localized impairment followed by complete or near-complete recovery. Generally, exacerbations recur over a period of years and disability accumulates. Within 10 years, approximately half of relapsing-remitting patients go on to a secondary progressive phase. The secondary progressive phase is characterized by gradual deterioration in neurological function that may be interspersed with acute relapses, followed by incomplete recovery. There is accumulation of neurological disability, leading to death in some patients. End-stage MS is characterized by paraplegia, ataxia, incontinence, and mental dysfunction. The period between the first symptom, often noted only retrospectively, and the more persistent and severe episodes may be several to ten years.

The prevalence of MS is lower in equatorial areas of the world than in the temperate latitudes, although this is more predominant in the northern hemisphere, where areas such as the northern US have prevalence of approximately 1 per 1000. Caucasians have higher incidence than other races, even at the same latitudes. Epidemiological studies indicate that the lifelong risk of disease onset is more closely related to the world region of life prior to mid-adolescence than to world region of residence after approximately age 15. The incidence of MS is approximately 2 to 3 times higher in women than men, thus 65 to 75% of patient populations in clinical practice and in trials are women. There are genetic factors in the incidence of MS as well. The best established of these is the HLA antigen DR2. HLA-DR2 positive individuals appear to have an increased incidence by a factor of 3 to 5. MS has a unimodal age-specific onset curve, with approximately 2/3 of cases occurring between the ages of 20 to 40 years old.

Diagnosis, especially for inclusion in clinical trials, has been codified over the years by consensus of the field, and published as formalized criteria and categories (Poser et.al, 1983). Diagnosis generally requires confirming at least two lesions which must have occurred in different parts of the CNS and at different times (demonstrating dissemination of disease activity in both time and space). Magnetic resonance imaging (MRI) has become a standard procedure in the diagnosis of MS. MRI readily demonstrates the MS lesions scattered throughout the brain. While the lesions are not pathonogmonic for MS, the pattern of lesions can be strongly suggestive.

Most patients have the relapsing-remitting form of MS at time of diagnosis, where the majority of symptoms are due to the acute episodes (usually termed attacks or exacerbations) which will usually resolve to a high degree. Over the ensuing years, as attacks involve overlapping portions of the CNS where prior attacks had left minor, often unnoticed residual deficit, the deficits accumulate. and exacerbations will only partially remit, with long-term accumulation of the incompletely resolved disability. Gradually, many patients shift into a form of the disease where there is increase of the disability without a clearly distinguished exacerbation having occurred. This is termed the progressive or chronic progressive form of the disease. The progressive form may have evolved from the relapsing-remitting form, in which case it is referred to as secondary chronic progressive. In a minority of MS patients, the progressive character predominates from onset in which case it is called primary progressive MS. There is of course a period of time in many patients when the relapsing-partially remitting character has not been entirely lost, but the progressive aspects are also prominent. This is often referred to as a relapsing-progressive stage.

The duration of the disease is highly variable. A small minority of patients have rapidly progressive disease and death within a few years of diagnosis, but most will have slowly progressive disability of several decades. The median duration of disease is greater than 30 years, with perhaps half of the deaths related to complications for which a predisposition was brought on by the disabilities. While the physical disability due to motor impairment is the deficit most focused upon, cognitive and neuropsychiatric disorders are common and often cause patient disability of equal magnitude to the motor impairment. Depression and suicide are increased in MS patients.

In the earlier stages, the clinical manifestations of disease activity are highly intermittent. This is in marked contrast to the current understanding of the pathophysiology of the disease, which has been revolutionized in recent years by the use of MRI. MRI demonstrates the presence of lesions on T2-weighted scans which many investigators believe represent fixed lesions due to temporally distant attacks on the brain. These lesions tend to develop slowly over time, and often show extensiveness that appears to have developed without concomitant clinical

symptomatology. This gave rise to the concept of "silent lesions;" MRI (and previously CT) lesions, especially in the cerebrum, that have developed without a corresponding history of clinical events. In addition to T2 lesions there are lesions visualized on T1-weighted MRI scans performed after iv infusion of a MR contrast agent, gadolinium currently the most popular. These gadolinium-enhancing lesions have a more fluctuating presence compared to the T2 visualized lesions. The gadolinium enhancing lesions vary over time, arising and resolving on a month-to-month time-scale. All of this MR evident disease activity may be present without any clinical manifestation. The meaning of these types of MR evident lesions, as relates to both improved understanding of the pathophysiology of the disease, as well as the clinical implications of the MR lesions, is a rapidly changing area which is a focus of research efforts.

Treatment of Multiple Sclerosis

MS therapies can be broadly divided into two categories: those directed against the immune system and intended to inhibit the disease process, and those intended to reduce symptoms. In general, the former have been less successful than the latter; however, it is immune modulator approaches that are looked to for major advances in effective therapy.

Symptomatic Therapies

Numerous agents have been used for symptomatic benefit in MS. These include amantadine and pemoline for treatment of fatigue, baclofen (a muscle relaxant and antispasmodic), tizanidine and benzodiazepines to treat spasticity, urologic antispasmodics for bladder dysfunction, and a number of agents for neuropsychologic impairment and pain management, including benzodiazepines, antidepressants and anticonvulsants. None of these agents retard the progress of the disease.

Immune Therapy

Immune therapy for MS is based upon the premise that an autoimmune process is involved in causing damage to the nervous system.

Immune suppressants

Corticosteroids (including ACTH) have long been used for treatment for acute exacerbations. These agents can decrease the peak severity and duration of the acute exacerbations, but do not appear to prevent the long term progression of disability.

Other immune-suppressants have been tested for treatment of MS; however, their limited benefit and potential for significant side effects have prevented any from being widely used. Studies with azathioprine, cyclophosphamide and low-dose methotrexate suggest modest benefits in MS compared to toxicities. Their use in the US is limited. Recent trials with the immune-suppressant ---------- appear promising for slowing progression of disease in chronic-progressive MS; however, its role in the clinical management of MS is not yet established.

Glatiramer acetate (COPAXONE®, Teva Marion Partners), formerly known as copolymer-1, is an immune modifier approved for reduction of the frequency of relapses in patients with Relapsing-Remitting MS. It is administered by subcutaneous administration.

Interferons

Interferons are cytokines capable of exerting multiple biologic effects through the expression of over 30 genes encoding proteins with anti-viral, anti-proliferative and immunomodulatory

functions. First described 40 years ago as potent anti-vial agents, their specific immunomodulatory activities are dependent upon the type of interferon and the particular biological system. Interferons have been broadly categorized into two classes: Type I and Type II. Type I IFNs are composed of the α-IFNs (of which there are many), and IFN-β. Alpha-IFNs are in clinical use for the treatment of a variety of malignancies and viral diseases. INF-β is encoded by only one gene. The Type II IFN is IFN-gamma (IFN-γ, also known as immune interferon), a cytokine produced primarily by natural killer cells and T-lymphocytes. Originally characterized based on its anti-viral activities, IFN-γ also exerts anti-proliferative, immunoregulatory and pro-inflammatory activities and is thus important in host defense mechanisms. There is no significant homology between IFN-γ and IFN-β or the various IFN-α family proteins. IFN-γ binds to specific cell-surface receptors with high-affinity binding sites.

The interferon signal transduction pathway involves the Jak-STAT mechanism. Jak (for Janus kinase) is a family of protein tyrosine kinases that includes Jak1, Jak2, Jak3 and Tyk2. Upon activation by an appropriate ligand, cell-surface receptors dimerize and bind two Jak protein kinases. The Jak-receptor complex undergoes phosphorylation and the multimeric complex catalyzes the phosphorylation of STAT transduction proteins (Signal Transducer and Activator of Transcription). The signal transduction pathway for IFN-α and IFN-β involves Jak 1 and Tyk2, and the receptor complexes recognize STAT1 and STAT2. Once phosphorylated, the complexes dimerize and complex with a DNA-binding protein to initiate transcription of early response genes. The IFNγ signal transduction pathway utilizes Jak1 and Jak2, and the complex recognizes STAT1. These complexes form homodimers, which do not require additional binding proteins for gene regulation.

The Interferon betas are now commonly used in the treatment of RRMS, where they are thought to inhibit viral replication and cell proliferation and enhance immunomodulatory activities such as phagocytosis.

Interferon β-1b

In 1993, the FDA licensed interferon β-1b (Betaseron®, Chiron, Inc.) for commercial sale in the US. Interferon β-1b can be classified as an immunoregulatory cytokine. Its mechanism of action in MS is incompletely defined, but it appears to involve reducing the numbers and activity of auto-reactive immune cells, reducing the movement of activated T-lymphocytes into the CNS, and influencing immune cells and cells of the CNS to produce anti-inflammatory cytokines. The Chiron Interferon β-1b product is a lyophilized protein product produced by recombinant DNA technology in *Escherichia coli*. The native gene was obtained from human fibroblasts. The product differs from the natural form by substitution of a serine for a cysteine residue at position 17, and by the ----------. Also in contrast to natural interferon β-1b, the product is not glycosylated due to its bacterial production method. The purified protein has 165 amino acids and an apparent molecular weight of 18,500 daltons. The specific activity of interferon β-1b is approximately 32 mIU/mg, based on a comparison of the product to the WHO reference standard of recombinant human interferon β.

Betaseron has been evaluated in >3000 patients in phase 1, 2 and 3 trials. A phase 3 trial in relapsing-remitting MS (RRMS), conducted between 1988 and 1993, demonstrated decreases in the relapse rate and the rate of accumulation of MRI lesions in the CNS. Based on that study, Betaseron was granted FDA approval for treatment of the relapsing-remitting form of MS. The main concerns regarding interferon β-1b have been its immunogenicity and it propensity to

cause injection site reactions and flu-like symptoms. There has also been concern in the field regarding depression, particularly early in the course of treatment.

An important unmet medical need in MS is the lack of an effective therapeutic agent for SPMS. Although interferon β-1b was shown to decrease the frequency of exacerbations in the RRMS study, there was no significant effect of Betaseron on disability progression. Thus, SPMS represents a sub-population of MS for which there is no specific approved therapy.

Two pivotal phase 3 trials of interferon β-1b for SPMS were subsequently initiated in the mid-1990s. The first (not under IND) was initiated in the EU in 1994, and terminated in February 1998 after a planned interim efficacy analysis showed statistically significant results. The data, summarized in a recent publication,[1] constitute the principal support for this supplemental BLA application and are analyzed in detail in this review.

[

]

Interferon β-1a

Interferon β-1a is produced in mammalian cells and receives the designation "1a" because its amino acid sequence is identical to that of the naturally occurring interferon β. In 1996, the first interferon β-1a for RRMS was licensed for commercial sale in the US (Biogen). The evidence in support of licensure was a phase 3 trial in which subjects with mild RRMS (EDSS 1.0-3.5) were treated with weekly 30 μg i.m. injections for one to two years. The trial succeeded on its primary endpoint of delay in EDSS progression, and there were statistically significant reductions in relapses, as well as both number and volume of MRI gadolinium-enhancing lesions. Despite the previous approval of Betaseron with orphan drug status, the Biogen product was granted orphan drug status because of data demonstrating that it was a different drug according to orphan drug regulations. This difference was based on the virtual absence of injection site necrosis typically associated with Betaseron use

An additional interferon β-1a (Rebif®, Ares-Serono Group) has been approved recently in Europe by the EC for RRMS and a BLA for Rebif® is under consideration in CBER. In the published randomized double-blind placebo-controlled PRISMS study,[2] 560 subjects in Europe and Canada with mild to moderate RRMS (EDSS 0 to 5.5) received s.c. interferon β-1a 22 μg, interferon β-1a 44 μg, or placebo on a TIW schedule for two years. Earlier studies had suggested a dose-effect with Rebif®; therefore a more intensive regimen was evaluated in this study (TIW rather than weekly dosing).

Overview of Prior and Ongoing Clinical Studies

Blinded, Controlled Studies

- [

---

[1] Placebo-controlled multicentre randomised trial of interferon β-1b in treatment of secondary progressive multiple sclerosis *Lancet* 1998; **352**, 1491-7

[2] PRISMS (Prevention of Relapses and Disability by Interferon β-1a Subcutaneously in Multiple Sclerosis) *Lancet* 1998; **352**, 1498-1504

]

- [

]

- Study TB01-35686/35886 consisted of two identical, multicenter, double-blind, placebo-controlled trials of interferon β-1b in RRMS. There were three parallel treatment groups (placebo, 1.6, and 8 mIU interferon β-1b [formally 9 and 45 mIU]). The study enrolled 338 subjects (226 subjects were randomized to interferon). Study treatments were self-administered s.c. on a QOD basis for approximately 2 years. This study provided the basis for licensure of interferon β-1b in RRMS. The studies were extended as protocols BL01-3103/3104.

- [

]

- Study 93079 is the pivotal study of interferon β-1b in SPMS. This study was conducted in 38 sites in 11 European countries. It provides the chief support for this sBLA and is reviewed in detail in this document.

- [

]

Open-Label Studies

- [

]

- [

]

- Study BL01-4114, the Betaseron patient experience study, was a phase 4 observational study that planned to enroll 1000 patients at 100 sites to quantify the incidence of predetermined adverse events. Due to slow enrollment, only 339 patients were enrolled at 27 sites. Forty-four patients (13%) discontinued interferon within the first 3 months, with two-thirds of these discontinuing because of the occurrence of one or more predetermined adverse events. One hundred twenty-three additional patients (36%) discontinued interferon between Month 3 and Month 24. Of these patients, half discontinued treatment because of the predetermined AEs, and half for other unspecified reasons. Overall, therefore, approximately 26% of patients discontinued interferon because of predetermined adverse events.

- [

]

- [

]

- [

]

Protocol ME 93079

Title: Double-Blind placebo-controlled multicenter study to evaluate the safety and efficacy of 8 million International Units Interferon β-1b given subcutaneously for up to three years to outpatients with secondary progressive multiple sclerosis

Study Period: September 1, 1994 to March 23, 1998 (last patient visit)

Funding: ----------

Centers: Thirty-eight sites within 32 European centers: Germany (7), United Kingdom (7), France (5), Italy (3), Netherlands (2), Finland (2), Ireland (1), Belgium (1), Spain (1), Sweden (1), Switzerland (1), and Austria (1). One German center was comprised of 5 sites; one French center was comprised of 3 sites.

Statistician: Lisa Bedell

Study Background

Berlex submitted two protocols for the evaluation of Betaseron in subjects with secondary progressive MS in mid-1995. A North American trial, protocol BL01-3112, is being conducted under IND ------- - a multicenter study designed to assess the safety and efficacy of two doses of Betaseron in SPMS. One active treatment arm is receiving a fixed dose of 0.25 mg Betaseron QOD; another active arm is receiving a size-adjusted dose of 0.16 mg/M^2 QOD; a third arm is receiving placebo. Protocol 93079 was submitted primarily as an informational amendment, and was intended to support registration of Interferon β-1b in Europe. It was not performed under formal IND regulations. The placebo-controlled study was designed to evaluate the safety and efficacy of Betaseron 8mIU SC QOD, and was to be conducted jointly by ---------- in Europe. The prior licensure of Betaseron for relapsing-remitting MS was on an exacerbation endpoint, whereas the primary endpoint for both studies was time to progression of disability.

In September 1995, CBER informed the sponsor that the results of Study 93079 would not be adequate to support expansion of the indication to secondary progressive MS. There was no response from ---------- to CBER until October 1997, however, by which time the study had progressed well beyond the North American study initiated in the same patient population. At that time, late in the conduct of the study and just prior to a planned interim efficacy analysis, the sponsor wished to address potential deficiencies in the study that might impair its ability to support, on a sole basis, expansion of labeling to include secondary progressive MS. Several areas of concern were identified and communicated to the sponsor in late-1997. The sponsor attempted to address these concerns through amendments to the statistical analysis plan.

Objectives

The stated primary study objective was to compare the efficacy, safety and tolerability of 8 mIU of Interferon β-1b to placebo in subjects with secondary progressive multiple sclerosis, when administered subcutaneously every other day for 36 months. The primary measure of efficacy was time to confirmed disease progression, as assessed by the Kurtzke Expanded Disability Status Scale (EDSS) score. Numerous secondary endpoints, delineated In the original study protocol, were reduced to four as a result of discussions with CBER in late 1997. The four selected secondary endpoints include: time to becoming wheelchair-bound, annual relapse rate, percent change in lesion volume and lesion activity as assessed by MRI.

Design

Overview

This was a double-blind, placebo-controlled study of two parallel treatment groups of outpatients with secondary progressive MS, performed at 32 European centers. Subjects were randomized to Interferon β-1b, 8 million International Units (mIU), or placebo, administered subcutaneously on a QOD schedule for 36 months. For the active treatment arm, Interferon β-1b was initiated at a dose of 4 mIU QOD for two weeks (7 doses), with a full 8 mIU QOD thereafter. The trial, as planned, consisted of a screening phase of up to 4 weeks, a treatment phase of 36 months, and

a drug-free follow-up period of 3 months. Study agents were self-administered or administered by caregivers. Subjects who terminated treatment prematurely were followed up at scheduled visits unless they withdrew consent or were lost to follow-up. The primary efficacy analyses were performed using the interim analysis dataset, including all data from all patient visits up through 11/20/97.

Randomization

Subjects were randomly assigned to receive either 8mIU Betaseron or placebo in a 1:1 ratio. Randomization was performed centrally by ---------- using a ------ program in advance of the study. Randomization was blocked within each site, with a fixed block size of six (6). Subject numbers were assigned in numerically ascending order at the time of enrollment. The randomization list was provided to the manufacturing department at Chiron, which was responsible for generation of vial labels and emergency envelopes.

Blinding

Study treatment vials were labeled by ---------- of Chiron Corp. Vial labels included the subject number, substance number, a code for the lot number, study number, quantity, mode of administration and expiration date. Labels for Betaseron and placebo were indistinguishable.

A number of measures were instituted to protect the integrity of the blind. Because side effects associated with Betaseron treatment are distinct and well known, the potential exists for physicians to surmise treatment assignment based on the presence or absence of side effects and laboratory abnormalities. The primary efficacy endpoint was derived from serial EDSS assessments; therefore, all routine EDSS evaluations were performed by designated EDSS Physicians, separate from the Treating Physicians. EDSS Physicians lacked access to clinical information and were prohibited from speaking to subjects except as necessary to perform standardized neurological evaluations. Results of neurological examinations, Functional Systems and EDSS evaluations were documented in a casebook which was maintained exclusively by EDSS Physicians. Manuals and training sessions were held for EDSS Physicians. Separate Treating Physicians were responsible for the overall medical care of subjects, as well as the evaluation, management and recording of exacerbations and adverse events. Treating Physicians were prohibited from communicating any information to EDSS Physicians that might lead to unblinding. Because cutaneous injection site reactions, in particular, are associated with Betaseron treatment, all potential injection sites were to be covered by standardized clothing during EDSS evaluations. Patients sensitive to changes in body temperature were instructed to take ibuprofen 400-600 mg orally TID concomitantly with the study medication for the first three (3) months, or throughout the study, if indicated.

Adequacy of blinding was evaluated with a blinding questionnaire, which was completed by Treating Physicians, EDSS Physicians and patients. The questionnaire was filled out at study completion, and by each rater only after the last patient had completed the study at their respective site. The three choices provided were: “placebo,” “Betaseron” and “don’t know.”

Reviewer's Comment:
An answer of “don’t know” enables the respondent to avoid making a best guess; more importantly, respondents who believe they know the treatment assignment may be tempted to respond “don’t know,” because they believe that a correct answer indicates unblinding and undermines the study, whereas a “don’t know” response suggests adequate blinding and strengthens the validity of the study. This potential bias in favor of providing a “don’t know” response limits the usefulness of the questionnaire.

Patient Population

The intended patient population was subjects with secondary progressive MS with moderate to severe disability, not wheelchair-bound at study entry. Planned study size was 720 subjects.

Inclusion Criteria

- a diagnosis of definite MS for at least one year
- MS in secondary progressive phase (history of relapsing-remitting disease followed by progressive deterioration sustained for $\geq$ 6 months)
- Evidence of clinically active disease (a history of $\geq$ 2 clearly identified relapses or deterioration of $\geq$ 1 EDSS point within previous 24 months; a 0.5 point increase was considered equivalent to a 1 point increase if baseline EDSS score was $\geq$ 6.0)
- No relapse or relapse-related neurological deterioration within 30 days prior to study entry
- Kurtzke Expanded Disability Status Scale score (EDSS score) between 3.0 and 6.5, inclusive (3.0 $\equiv$ moderate disability in one functional system, fully ambulatory; 6.5 $\equiv$ constant bilateral assistance [canes, crutches, braces] required to walk $\approx$20 meters without resting)
- male and female subjects, age 18-55 inclusive

Exclusion Criteria

- any form of MS other than secondary progressive
- any other disabling condition that could interfere with the clinical or MRI evaluation
- pregnancy or lactation
- alcohol or drug abuse in the 90 days preceding screening visit
- uncontrolled clinically significant heart disease (angina pectoris, congestive heart failure, dysrhythmias)
- clinically significant hepatic dysfunction (SGOT > 3X upper limit of normal range)
- clinically significant renal dysfunction (creatinine > 180 μmol/L)
- clinically significant bone marrow dysfunction (hemoglobin < 8.5 g/dL, WBC < 2.5 X 10^9/L, or platelet count <125 X 10^9/L)
- any prior use of: total lymphoid irradiation, interferons, other recombinant DNA cytokines, murine antibodies or any T-cell antibody
- within 24 months: cyclophosphamide, mitoxantrone, other immunosuppressive therapy or cytotoxic chemotherapy
- within 12 months: 15-deoxyspergualine, other immunomodulatory drugs, azathioprine (frequent MRI-subgroup)
- within 6 months: cyclosporine A, azathioprine (all other patients), IgG
- within 1 month: corticosteroids, ACTH
- history of suicide attempt or current suicidal thoughts

Written informed consent was obtained from all participating subjects.

Treatment

Material Source

Interferon β-1b was produced for ---------- by Chiron Corporation, Emeryville, CA. Each vial of Betaseron contained 0.3 mg (9.6 mIU) Interferon β-1b as a lyophilized powder, which was reconstituted with 1.2 mL 0.54% sodium chloride prior to injection. Placebo vials contained human albumin and D-glucose only. Labeling of the vials was performed by the ---------- --------

----------of Chiron Corporation, ----------. There were no changes in the appearance of the study drug or packaging during the study, except that the study agent was sealed in 5 mL vials for the first 6 months of the study, and 3 mL vials thereafter.
An EU Commission decision (dated July 30, 1997) requested elimination of the use of specific materials of bovine and ovine origin in the production of biological products. In accordance with this decision, the study medication used in 1998 was manufactured using a new production process which involved a change in a growth medium used in an early production stage of Betaseron. The drug product used in the study was the same drug product commercially available in the U.S. (manufactured by Chiron).

Dose and Administration

One milliliter (1.0 mL, 8 mIU) of the reconstituted interferon β-1b solution or an equivalent volume of placebo was administered every other day by subcutaneous injection by the subjects or their caregivers. This represents the interferon β-1b dose presently approved for relapsing-remitting MS. Treatment compliance was monitored through a drug accountability procedure. Subjects had the option to record days and doses of study drug administration in a diary. Treatment was discontinued for grade ≥ 3 toxicity, and could be reinstituted at 50% dose once toxicity had fallen to grade ≤ 2. Full dosing could resume after 2-4 weeks. If a full dose of study medication could not be tolerated, the Treating Physician could maintain treatment at a lower dose, but not less than one half the normal dose. Patients who discontinued treatment but underwent scheduled evaluations could recommence treatment.

Concomitant Medications

Recommended: Ibuprofen (400-600 mg TID) was recommended (but not mandated) for prophylaxis against fever, myalgias or other flu-like symptoms during the first 3 months of study agent dosing. If not tolerated, paracetamol (500 mg QID) or indomethacin could be given. These medications were to be administered simultaneously with study drug injection, and could be continued for the duration of the study, if indicated. If indicated, ranitidine (150 mg) was given nightly for the first 2 weeks of study treatment.

Specially Directed: Systemic glucocorticoids could be administered according to two protocol-specified treatment schedules (schedule 1: IV infusions of methylprednisolone 1000 mg X 3 daily doses, followed by oral prednisone / prednisolone on a 15 day tapering course; schedule 2: IV infusions of methylprednisolone 1000 mg X 3 daily doses, only). Courses were to be limited to three (3) courses or fewer within 12 study months, if possible.

Prohibited:

- immunomodulatory or immunosuppressive treatments
- other therapeutic agents for MS
- other investigational therapy for MS
- systemic steroids, other than that allowed for treatment of exacerbations

Reviewer's Comment: Use of ibuprofen and paracetamol was not uniform across all sites. This could result in inconsistency between or within sites, as well as increase the potential for unblinding at sites not using prophylactic treatment. Such unblinding could introduce bias in subsequent patient reporting and evaluations.

Evaluations

Schedule of Evaluations

- Day 1, 3, 5, and 15; monthly X 3, then every 3 months:
  vital signs (including temperature), relapse assessment, Adverse Events
- Monthly X 3, then every 3 months:
  Treating Physician: physical examination
  clinical laboratories (hematology, blood chemistries, urinalysis, neutralizing antibodies)
- Every 3 months:
  EDSS Physician:
    neurological evaluation
    Kurtzke Functional Systems (FS)
    Ambulation Index (AI)[3]
    Kurtzke EDSS score (disability score)[4]
  Montgomery-Asberg Depression Scale (MADRS)[5]
- Every 6 months:
  Quality of Life (Sickness Impact Profile)[6]
  lipid profile, thyroid function studies
- Every 12 months:
  Magnetic Resonance Imaging (MRI) - T2 imaging
  Cognitive Test battery (476 patient subgroup)[7]
- At 36 months:
  Global Evaluation of MS (GEMS)

Frequent Magnetic Resonance Imaging (MRI) Subgroup:

- day 1, then monthly, months 1-6; monthly, months 18-24
  MRI before and after gadolinium contrast (T2 and T1)

Removal of Patients from Treatment or Assessment

Treatment could be discontinued at any time at the discretion of investigators. Treatment had to be discontinued in the event of: patient decision to discontinue treatment, intolerable Adverse Event(s), loss to follow-up, pregnancy, unauthorized use of study medication, use of other investigational/experimental therapies or chemotherapeutic agents for MS, break of the blind, suicidal ideation/attempt.

Efficacy Endpoints

Primary Endpoint

The prospectively-defined primary endpoint was time to neurological deterioration, defined as a one point increase on the EDSS scale from baseline. For subjects with a baseline EDSS of 6 or greater, a 0.5-point increase was considered equivalent to a 1-point increase. The EDSS is a

---

[3] the AI uses a 10 point scale; 0 = unrestricted; 9 = wheelchair bound and unable to transfer self independently

[4] 10 point scale in 0.5 point increments; 0=normal neurological exam; 7=essentially restricted to wheelchair; 10=death due to MS

[5] observer rating scale for psychological symptoms

[6] questionnaire with 136 items

[7] Rao Repeatable Battery of Neuropsychological tests

10 point scale in 0.5 point increments from 0 (normal neurological exam) to 10 (death due to MS). EDSS steps through 4.5 indicate no impairment in ambulation; steps 5.0 to 9.5 represents progressive impairment in ambulation and mobility. The EDSS scores were based on standardized neurological evaluations performed by EDSS Physicians. Data from the neurological evaluations were used to determine Kurtzke Functional Systems (FS) assessments. The FS scores assess function within individual neurological systems, including visual, pyramidal, cerebellar, brainstem, sensory, bowel and bladder, cerebral and 'other.' The final EDSS scores were based on the combination of FS scores and the subjects' level of mobility. The 'other' portion of the FS assessments was not utilized in the determination of EDSS in this study.

To be considered valid, EDSS increases had to be maintained and confirmed at one subsequent and consecutive scheduled study visit at least 70 days later, and maintained at any intercurrent visits. EDSS scores obtained during intervals categorized as exacerbations/relapses were not included in the EDSS analyses, and all scores obtained at unscheduled visits were omitted from calculation of the endpoint. The intent was to insure that transient relapse-related increases in EDSS would not contribute to the endpoint. Relapses for which a date of resolution was not recorded were considered to have a duration of 180 days, after which the EDSS scores were applicable to the endpoint. Missing EDSS scores were interpolated as the lower of the two bracketing values, thus assuring that a missing score could not be confirmatory of an increase in EDSS.

For calculation of the primary endpoint, the EDSS score obtained at the baseline visit was used as the initial EDSS; the score obtained at the screening visit was not used. EDSS scores recorded at Months 3, 6, 9, and every three months thereafter were used in the assessment of the primary endpoint. Because subjects went off medication after Month 36, the Month 39 visit was declared to be invalid for confirmation of an increase in EDSS first observed at the Month 36 visit. Thus, the final opportunity for progression to occur was the Month 33 visit.

Secondary Endpoints

The original analysis plan delineated numerous secondary endpoints to be evaluated using multiple statistical tests, with little regard to clinical meaningfulness or relative importance. CBER communicated this concern to the sponsor in late 1997, as well as a concern regarding increased Type I error associated with multiple comparisons. As a result of discussions with the Agency, the sponsor limited the number of secondary endpoints to four (4):

- Time to becoming wheelchair-bound
  the number of days from initiation of dosing to the onset of the first EDSS score of 7.0; confirmation was not required.
- Annual relapse rate
  the number of relapses, confirmed by investigators, divided by time-on-study.
- percentage change in T2 lesion volume from baseline to last scan available
  only subjects with a valid baseline scan and at least one valid on-study scan included in analyses
- number of newly active lesions during months 1-6
  lesions present on T2-weighted or gadolinium enhancing images which, relative to month 0, displayed new enhancement, were non-enhancing but new on the T2 scan, or were non-enhancing but showed enlargement on the T2 scan. Subjects with missing baseline scans were non included in the analyses.

Tertiary Endpoints

- EDSS Endpoints
    - proportion of subjects with confirmed disease progression
    - change in EDSS from baseline
    - EDSS at endpoint

      The change in EDSS was analyzed in terms of the endpoint EDSS (last EDSS available) with stratification by baseline EDSS. The evaluation of the change in EDSS was assessed both as actual change, and as change categorized as <0.5, 0.5, 1, 1.5, or ≥2.
- MRI Endpoints
    - absolute change in T2 volume
    - number of new or enlarged lesions as seen in annual MRI scans
    - proportion of active annual MRI scans (scans containing new or enlarged lesions)
    - proportion of patients with active annual MRI scans
    - During months 1-6 AND during months 19-24 (frequent MRI cohort):
        - number or persistently enhancing lesions as seen on gadolinium-enhanced monthly MRI scans
        - proportion of active scans
        - proportion of patients with active scans
    - number of newly active lesions during months 19-24 (relative to month 18) as seen on T2-weighted and gadolinium-enhanced monthly MRI

- Relapse Endpoints
    - time to first relapse
    - distribution of relapse severity
    - proportion of patients with moderate or severe relapses
    - annual rate of moderate and severe relapses
    - mean duration of relapses per patient
    - number of days per patient spent in a relapse
    - number of days per patient spent in a moderate or severe relapse
- Cognitive function endpoints (change from baseline for the following test variables):
    - selective reminding test
    - 10/36 spatial recall test
    - paced auditory serial addition test (PASAT)
    - symbol digit modalities test
    - word list generation
    - Rao Battery/ composite score
- Quality of Life: Sickness Impact Profile (SIP) Variables - used to assess sickness-related dysfunction in 12 areas of activity: sleep and rest; emotional behavior; body care and movement; home management; mobility; social interaction; ambulation; alertness behavior; communication; work, recreation and pastimes; and eating
    - change in sum of 12 individual scores, each expressed as percent of maximum dysfunction
    - change in physical dimension score (body care and movement, ambulation, mobility)
    - change in psychosocial dimension (emotional behavior, affective behavior, social interaction, and communication
    - overall SIP score as percent of maximum dysfunction
- Ambulation Index (AI)

    - change in AI from baseline
    - time to deterioration of ≥ 2 steps of the AI
    - proportion of subjects deteriorating by ≥ 2 steps
- Global Evaluation of MS (GEMS)
    - change from baseline to month 36
- Steroid Use
    - proportion of patients with steroid use
    - number of steroid courses per patient
- Efficacy results by study year
    - number of relapses during each year
    - change in EDSS during each year
    - change in lesion volume during each year
- Exploratory Analyses
    - enhancing lesion load in Months 1-6 and Months 19-24

Correlation Between MRI and Clinical Outcome

Prompted by communications between the sponsor and CBER in late 1997, an analysis was planned to evaluate the correlation of MRI with clinical outcome as part of the sponsor's phase 4 commitments from the 1993 marketing approval.

The primary *correlations are:*

EDSS

- change from baseline EDSS **and** percent change in T2 lesion volume for the time point corresponding to the last MRI performed (all subjects)

Relapse Rate

- annual relapse rate (2 year data, frequent MRI subgroup) **and** number of newly active lesions during month 1-6 and month 19-24 (using average ranks from both periods)

The secondary correlations are:

EDSS

- logrank scores for time to confirmed progression **and** percentage change in T2 lesion volume at the time point corresponding to the last MRI performed (all patients)
- change from baseline EDSS **and** number of newly active lesions during month 1-6 and month 19-24 (average ranks from both periods)
- change from baseline EDSS **and** percent change in T2 lesion volume during year 1

Relapse Rate

- annual relapse rate **and** number of newly active lesions during month 1-6 (frequent MRI subgroup)
- annual relapse rate **and** number of active scans (scans with newly active lesions; frequent MRI subgroup)
- annual relapse rate **and** percent change in T2 lesion volume
- annual relapse rate **and** percent change in T2 lesion volume during year 1

In addition, rank correlation coefficients were to be calculated for T2 lesion volume and neuropsychologic assessments, EDSS scores and QOL measures, MRI lesion activity and MRI lesion area.

Additional exploratory multivariate analyses was planned to assess the relation between EDSS changes and MRI variables.

Safety Endpoints

The safety population was to include all patients who received at least one administration of the study treatment and who provided at least some post-baseline safety data. Safety endpoints were to include the following assessments:

- Adverse Events - were graded as: 1) mild - if the subject was aware of symptoms or signs, but the symptoms or signs were easily tolerated; 2) moderate - if symptoms or signs were sufficient to restrict but not prevent usual daily activity; or 3) severe - if the subject was unable to perform usual daily activity. The proportions of subjects with each event were to be presented for each treatment group. Fisher's exact test (two-sided p-values) would be used for statistical comparisons, with presentation of two-sided 95% confidence intervals for events which occurred in $\geq$ 5 patients in the interferon β-1b group. Additional analyses for flu-like symptoms and injection site reactions were to be analyzed through life-table estimates (for analysis of time to first occurrence), as well as frequency and severity. These analysis would be performed in an attempt to clarify the timing of these events (first few weeks of treatment versus later).

- Laboratory Variables - Descriptive statistics were to be presented for continuous laboratory variables per treatment at baseline and for changes from baseline at subsequent time points. Individual patient changes were to be evaluated with shift tables. Proportions of patients with new or worsened abnormalities were to be compared between treatment arms using Fisher's exact test at each time point. Laboratory variables which are graded on the clinical toxicity scale were to be evaluated in frequency tables. Comparisons were to be performed using the Fisher's exact test with respect to categories defined by grade < 2 versus grade $\geq$ 2.

- Vital Signs - Descriptive statistics were to be presented per treatment for actual values at baseline and for changes from baseline at subsequent time points. Changes were to be compared between groups using two-sided tests; 95% confidence intervals were to be presented regarding changes within treatment groups. Fisher's exact test was to be used to compare rates of abnormally high blood pressure ($\geq$ 150/100 mmHg) between groups.

- Electrocardiograms (ECG) - The proportions of patients with new ECG abnormalities were to be compared between groups using Fisher's exact test.

- Physical Examinations - Patients with new abnormalities on physical examination were to be evaluated descriptively.

- Concomitant Medications - Concomitant medication use was to be summarized descriptively, with emphasis on the effects of NSAID and paracetamol on the rate of occurrence of flu-like symptoms and fever.

- Serum Neutralizing Antibodies (NAB) - NAB activity was to be assessed at baseline, months 1, 2 and 3, and every 3 months thereafter. The effect of NAB activity on safety and efficacy was to be evaluated. Subjects were considered to be NAB positive if serum titers were $\geq$ 1:20 on two consecutive visits, with time to NAB positivity the interval from baseline to the first of the two visits. Exploratory analyses were performed using additional titer cutoffs of 1:100 and 1:400.

- Montgomery and Asberg Depression Rating Scale (MADRS) - A tool to assess depression, the MADRS is based upon queries of psychological symptoms in 10 areas (each on a 0 to 6-point scale). Originally, the MADRS was to be assessed at baseline, 3 and 6 months, and every 6 months thereafter. In October 1995, the protocol was amended to assess MADRS every three months. MADRS scores were to be presented for each time point with descriptive statistics. As outlined in the final (February 1998) protocol amendment, the scores were to be categorized as "normal," "mild," "moderate" or "severe" based upon point total. Changes in MADRS score were to be compared between groups using the Wilcoxon test. The numbers of categories shifted (i.e., normal→moderate = 2) were to be assessed using the Wilcoxon test. Analyses of proportions of patients were to be compared using Fisher's exact test.

Reviewer's Comment: The MADRS has not been used in MS prior to this study and is not properly validated. The test is designed to be used to assess *changes* in psychological symptoms. The MADRS was not administered until month 3, thus early and transient depression would not be captured by this study. Moreover, the analysis of the Month 3 scores are entirely dependent on the Month 0 scores (because *change* is assessed), and Month 0 scores may be less reliable than subsequent scores because of lack of familiarity on the part of both patients and test administrators. The test should be administered at the same time each day; however, this was not a requirement of the protocol.

Interim Analyses

An interim efficacy analysis was prospectively planned after all subjects had completed 24 months on study. In the event of early study termination, the dataset at the interim cut-off date was to be the primary dataset for the primary, secondary and some of the tertiary variables. In the event of important inconsistencies, the data collected between cut-off date and termination were to be explored further. The database for the Advisory Board interim analysis was locked on November 20, 1997, at which time all subjects had completed at least 2 years on study. The study was terminated by the Advisory Board at its meeting on January 18, 1998 due to favorable results.

Planned Final Analyses

Primary Endpoint

Time to Confirmed Progression

The primary efficacy variable was analyzed using a Cochran-Mantel-Haenszel (CMH) test for covariance-adjusted logrank scores. Covariance was adjusted for baseline EDSS as a categorical variable, with values of 3 and 3.5 combined because of the small sample sizes. There was also a stratification adjustment for center. Multiple supportive statistical methods were to include: Mantel-Cox logrank test stratified for baseline EDSS ($\leq$3.5, 4-5.5, $\geq$6) and time interval, Mantel-Cox logrank test stratified for time interval and study center, and piecewise logistic regression modeling. Other factors, including duration of MS, age, gender, body surface area, and baseline T2 lesion volume, were evaluated in the logistic regression model, as well as their interaction with treatment or time interval. The primary analysis of time to confirmed progression was an intent-to-treat analysis which included all data for all subjects, both during and after study treatment, regardless of whether subjects went on to receive active treatment for MS.

Handling of Missing Data, Protocol Deviations

The sponsor performed supportive analyses to deal with missing data and protocol deviations. These are briefly summarized below:

- An analysis of a modified intent-to-treat population included all subjects who had used study treatment for ≥6 months and undergone planned EDSS evaluations through that period. Subjects were analyzed in treatment groups based on actual treatment received.
- An analysis of an efficacy evaluable population included only subjects who completed the first year on study without major protocol violations.
- An intent-to-treat "A" analysis considered subjects lost to follow-up as having confirmed progression during the 3 months following the loss.
- An intent-to-treat "B" analysis considered subjects lost to follow-up as not having confirmed progression during the 3 months following the loss.

Secondary Endpoints

Time to Becoming Wheelchair Bound

Treatment groups were compared using Mantel-Cox test stratified for baseline EDSS, with additional analyses based on piecewise regression models as described for time to progression.

Annual Relapse Rate

Treatment group differences in annual relapse rates were assessed using a non-parametric analysis of covariance with stratification adjustment for center. Standardized ranks of relapse rates, assigned across centers, were fit to a linear model which included a dichotomous indicator for relapses in the 2 years prior to enrollment and center (no relapses; ≥1 relapse). Residuals from the linear model were used to evaluate treatment group differences using a CMH test stratified for center.

Percentage Change in Lesion Volume

Absolute and percentage changes in lesion volume (baseline to last scan) were compared using a non-parametric analysis based on a CMH test, stratified for center, as described for the analysis of annual relapse rate (above).

Number of Newly Active Lesions During Months 1-6

This analysis was also performed with a non-parametric analysis of covariance. The number of newly active lesions and baseline lesions were converted to standardized ranks within each center, and the analysis was performed using a CMH test as described for annual relapse rate (above).

Tertiary Endpoints

Numerous supportive efficacy analyses were performed using CMH tests including proportion of subjects with confirmed progression, EDSS at end of study, change in EDSS from baseline, MRI lesion activity, relapse rates, time to first relapse, relapse duration, cognitive functions, quality of life assessed by Sickness Impact Profile, ambulation index, Global evaluation of MS, steroid use and correlations between MRI and clinical outcome.

Significance Testing, Allocation of Alpha

The prospectively defined criterion for statistical significance for the interim efficacy analysis was $\alpha = 0.0133$ for a two-sided test. All secondary and tertiary endpoints were considered to be significant at $\alpha = 0.05$. No adjustment for multiple analyses was planned, because the success of the trial was to be determined by a single analysis (time to confirmed progression).

Study Performance

Study Administration

The study commenced in September 1994, well before first submission of the protocol to CBER. The trial was guided and supervised by the Study Director from ----------. The Study Director also chaired the Study Steering Committee, which provided advice and recommendations with regard to study conduct. The Steering Committee remained blinded to treatment assignments throughout the study. A Study Advisory Board, comprised of experts in neurology or biostatistics who were not directly participating in the conduct of the study, monitored safety data and performed interim safety reviews at intervals of one year or less. The Advisory Board was responsible for notification of the study sponsor and Steering Committee of relevant safety issues that became apparent during review of the data. Assay for neutralizing antibodies was performed centrally by ----------. All other laboratory samples were analyzed centrally by ---------- A contract research organization (----------) entered data from the case report forms, and developed the majority of the programs for statistical analyses and tabular displays. Additional analyses were performed by ----------.

In November, 1997, in preparation for its January, 1998 Advisory Board Meeting, the sponsor initiated discussions with CBER regarding potential deficiencies in the study design and analytic plan. ----------attempted to address these concerns through amendments to the statistical analysis plan.

As a result of the its meeting on January 18, 1998, the Advisory Board recommended to the study sponsor that the study be terminated, and the investigators were notified of study termination on February 9, 1998.

Formal Protocol Modifications

The protocol underwent five revisions, summarized below. With the exception of amendment 5, protocol amendments were not submitted to CBER at the times of modification.

Amendment 1 (2/21/95) added a Cognitive Function subgroup, more specific guidelines for prophylactic use of NSAIDs, increased the number of trained EDSS Physicians at each site from one to two, and outlined multiple minor administrative and procedural changes.

Amendment 2 (10/20/95) added a provision for Treating Physicians to lower the dose of the study agent to less than full dose (but at least one half of a full dose) in the event of toxicity. The statistical section (section 9) was deleted and reincorporated as Appendix N.

Amendment 3 (6/5/97) added a post-study blinding questionnaire and increased the frequent MRI Subgroup from 108 patients at 6 centers to 125 patients at 7 centers. In addition, multiple secondary efficacy variables were reclassified as either secondary or tertiary variables.

Amendment 4 (9/10/97) changed the criteria for early termination of the study on the basis of the interim efficacy analysis. The primary analyses of confirmed progression were restricted to the conservative intent-to-treat and modified intent-to-treat populations. The analysis for the efficacy evaluable population was removed from the primary endpoint.

Amendment 5 was submitted to IND---------- in December 1997 following considerable communications and discussions with CBER. This amendment substantially revised the statistical analysis plan and stipulated that the revised plan would be used to guide the Advisory Panel regarding a decision to stop the trial. The amendment also directed that all statistical analyses would be performed by ----------, rather than the ---------- Biometrical Department.

Patient Enrollment

The first subject was enrolled at the Berlin site on September 5, 1994; enrollment commenced at all sites within an 11-month period, with the last patient enrolled on August 3, 1995. The final patient visit occurred on March 23, 1998. Seven hundred sixty-eight (768) subjects were screened; 718 were entered into the study and randomized to treatment. All 718 randomized subjects received at least one dose of the study agent and comprise the ITT population of the study. The distributions of patients between treatment arms, sites and countries are summarized in Table 1:

Table 1: Breakdown of Numbers of Subjects Enrolled by Site

Country	Study Site	Placebo	Interferon	overall
Germany	Berlin	30	30	60
	Osanbrück	9	9	18
	Würzberg	10	9	19
	Munich	7	9	16
	Erfurt	11	10	21
	Göttingen	9	9	18
	Düsseldorf	6	6	12
Switzerland	Basal	9	9	18
Austria	Vienna	6	6	12
UK	Birmingham	9	9	18
	Cardiff	15	15	30
	Newcastle-upon-Tyne	6	6	12
	Belfast	12	12	24
	Sheffield	9	9	18
	Aberdeen	12	12	24
	London	12	12	24
	Dublin	15	15	30
France	Rennes	9	9	18
	Bordeaux	9	9	18
	Toulouse	12	12	24
	Paris	12	12	24
	Lyon	9	12	21
Italy	Milan	12	12	24
	Florence	9	9	18
	Rome	15	15	30
The Netherlands	Amsterdam	15	15	30
	Groningen	18	17	35
Belgium	Melsbroek	12	12	24
Sweden	Huddige	9	9	18
Spain	Barcelona	12	12	24
Finland	Turku	9	9	18
	Helsinki	9	9	18
Total		**358**	**360**	**718**

Protocol Deviations

Protocol deviations were reported in 27% of placebo subjects and 20% of Betaseron subjects and are summarized in Table 2. Eligibility violations were reported in 56 of the 718 subjects. All but one violation was for lack of clinically active disease (clinically active disease defined as a history of at least 2 clearly identified relapses or progression in EDSS within the 24 months prior to enrollment). One subject violated exclusion criteria on the basis of use of a prohibited substance prior to the study. It is notable that control subjects had approximately twice as many steroid and concomitant medication use deviations as Betaseron-treated subjects. The sponsor's explanation for these disparities was a greater frequency of active disease in placebo subjects.

Reviewer's comments on sponsor-identified deviations:

- The deviation "no evidence of clinically active disease" refers to subjects with neither confirmed EDSS progression nor 2 exacerbations in the 24-month pre-study period. Such subjects can not be appropriately classified as having SPMS, and do not belong to the target patient population. This is a potentially serious infraction; however, because the numbers of subjects in this category are balanced

Table 2: Protocol Deviations

Deviation	Placebo n (%)	Betaseron n (%)
Eligibility Violations		
No evidence of clinically active disease	26 (7.3)	30 (8.3)
pre-study use of prohibited substance	1 (0.3)	
On study Violations		
> 3 courses steroids per year	20 (5.6)	14 (3.9)
prohibited concomitant medications	45 (12.6)	19 (5.3)
EDSS not performed by EDSS physician	1 (0.3)	6 (1.7)
non-standard clothing for neuro exam	2 (0.6)	2 (0.6)
Total	**95 (26.5)**	**71 (19.7)**

between the two treatment arms (approximately 8% in each), these patients should not affect the study endpoints directionally.

- Greater use of steroids and prohibited concomitant medications in the control group is consistent with the concept that placebo-treated subjects had generally more MS-associated symptoms during the study. The greater use of ancillary medications would tend to decrease the severity of exacerbations, or to reduce their apparent rate. Thus, excess use in the control group has the potential to decrease the apparent severity of disease relative to the active treatment arm, which would strengthen the conclusions regarding a Betaseron treatment effect.
- Interferons have salutary effects in MS. Eight subjects in the interferon group discontinued blinded treatment in favor of open interferon. Seven of these subjects were treated with interferon β-1b, one received interferon β-1a. Fifteen (15) subjects in the placebo group discontinued blinded treatment in favor of open interferon β-1b (13 subjects) and interferon β-1a (2 subjects). The observed excess use of interferons in the placebo group would tend to diminish symptoms to a greater extent in that group, thereby strengthening the conclusions regarding a Betaseron treatment effect.

Reviewer's comments on CBER-identified protocol violations:

- One patient was older than the protocol-specified 55 year limit.
- Two subjects were enrolled with baseline EDSS <3; two subjects were enrolled with EDSS > 6.5.
- Physicians were to try to limit subjects to no more that three courses of steroids per year of study. Whereas the sponsor reported 20 and 14 violations in placebo- and Betaseron-treated subjects, respectively, violations were found in 27 and 23 subjects in these respective groups upon review of the data. This difference probably relates to the definition of a "year of study," i.e., a calendar year, versus a 12-month period synchronized with the date of first administration of the test agent, versus any contiguous 12-month period. This disparity was found using the latter definition of a contiguous 12 month period. This difference is not deemed to be important, because the violations for frequent steroid use are directionally the same as reported by the sponsor and more frequent violations in the placebo group tend to strengthen the sponsor's conclusions.

- All subjects were to have completed Baseline Visit evaluations within a 10 day period, and were to receive initial administration of the study agent within this interval. In fact, only 229 of 718 subjects received their initial dose of study agent within the specified 10-day time-frame (Table 3). The mean time from screening to first study agent was similar in the two groups (16.2 days, interferon; 16.0 days, placebo). In instances when an exacerbation occurred between screening and the first dose of the study agent, subjects were to undergo additional screening and baseline evaluations, which would serve to lengthen this interval. There is no information in the submission with regard to these subjects, and such data are not thought to be materially important for this review because it is likely that this occurred in only a small fraction of subjects.

Table 3: **Time From Screening to First Study Agent Administration**

Interval	Frequency
within 10 days, per protocol	229
10-20 days	263
21-30 days	140
31-40 days	61
41-50 days	20
51-60 days	2
61-70 days	2
>70 days	1

- EDSS evaluations between study month 3 and study month 39 were to be completed within 10 days of schedule. For visits between 3 and 33 months, inclusive, 70% were performed within the protocol-specified 10-day time limit; 83% and 99% were performed with 2 weeks and one month of the target date, respectively. Month 36 and month 39 visits were not included in this analysis because many of the subjects who remained on study at the time of trial termination had their final EDSS evaluations incorrectly entered as month 36 visits, when in fact they occurred much sooner. It is noteworthy, however, that true month 36 EDSS evaluations were performed late in only 7 subjects (10 to 68 days after scheduled date).

Randomization

Adherence to the blocked randomization scheme was maintained such that balance of treatment assignment within sites was excellent. Using a block size of 6, the theoretical maximum imbalance between treatment groups within sites is 3; only one site (Lyon, France) exhibited imbalance of this magnitude.

Time-On-Study

Time-on-study (enrollment to last date eligible to report exacerbations or adverse events) represents the duration of opportunity for exacerbations or adverse events to occur. For the complete data set, time-on-study ranged from 0 to 39 months, with a median of 35.4 months. Eighty-seven percent of subjects completed 33 weeks or more on study, and 67% completed 36 weeks. Given the requirement for confirmation of progression on a second 3 month evaluation, the penultimate scheduled EDSS evaluation provided the final opportunity for progression (time at-risk). The mean times at-risk were 31.5 and 31.1 months for interferon and control groups, respectively (31.3 months, overall). Ten subjects (5 per group) participated for 5 months or less, such that confirmed progression was not possible.

Reviewer's Comment: There was no prospective rule for subjects lost to follow-up before Month 5 (i.e., censoring at time zero, time of loss, or Month 5); however, CBER's examination of the sponsor's data sets indicates that these subjects were censored at the time of loss (as were all subjects lost to follow-up). Although the lack of a prospective plan was an oversight, the time of censoring of these 10 subjects within the initial 5 months of the study could not importantly impact the study results.

Time-On-Treatment

Overall, 31% of subjects discontinued the study agent prematurely. Throughout the 36 month study period, discontinuations were evenly distributed between treatment groups by study period. Discontinuations were relatively constant in both groups with respect to time, approximately 10% per year of treatment in both groups.

The reasons for premature discontinuation of treatment differed in incidence between groups (Table 4). Relative to the Betaseron group, twice as many subjects in the placebo group withdrew their consent, and twice as many placebo subjects discontinued the study agent because of progressive disease. In contrast, twice as many Betaseron subjects prematurely discontinued the study agent because of adverse events relative to placebo subjects.

Table 4: Reasons for End-of-Treatment	Placebo (*N*=358)		Interferon (*N*=360)	
	N	(%)	*N*	(%)
completed study	239	(66.8)	256	(71.1)
Adverse Event	27	(7.5)	52	(14.4)
withdrawal of consent	29	(8.1)	11	(3.1)
progressive disease	56	(15.6)	27	(7.5)
protocol deviation	4	(1.1)	7	(1.9)
death			2	(0.6)
loss to follow-up	3	(0.8)	3	(0.8)
laboratory deviation			1	(0.3)
pregnancy			1	(0.3)

Reviewer's Comment: The sponsor provided (by time and treatment group) the numbers of subjects who discontinued study drug because of adverse events. CBER plotted these data for interferon subjects, and observed a linear relation between cumulative discontinuations and time. These data suggest that: 1) adverse events, sufficiently severe to warrant discontinuation of treatment, occur at a constant rate of approximately 5% per year; and 2) tolerability of interferon in the short- or intermediate-term does not predict tolerability in the long term.

Patient EDSS Evaluations

Patients were scheduled for evaluations on Day 1, 3, 5, and 15, monthly X 3, and then every three months during the course of the study. A total of 10,386 EDSS scores were recorded, including all screening and baseline scores and all scheduled and unscheduled visits. Following the screening and baseline evaluations, there were 9055 EDSS evaluations, of which 686 (7.6%) were Unscheduled Visits. The percentages of Unscheduled Visits were 8.6% and 6.6% in the interferon and placebo groups, respectively. Between months 3 and 39, inclusive, 8396 EDSS scores were recorded at scheduled visits. There were 520 scores obtained during exacerbations that were disallowed: 233 were in the active treatment arm and 287 were in the control arm.

Missing EDSS Scores

Values of missing scores were assigned as the lower of the two bracketing scores. Thus, missing scores could confirm progression only when the next available score provided such confirmation.

Reviewer's Comment: In a review of the SAS data files, CBER found no screening EDSS score recorded in 16 subjects, and no baseline score recorded in 79 subjects. In these latter cases, the screening EDSS was used as the initial EDSS score for calculation of the primary endpoint. During the study (beyond baseline), CBER found a total of 68 missing EDSS scores requiring interpolation (<1% of the total).

Treatment Accountability

The numbers of vials dispensed, and numbers of vials returned (opened and unused) were recorded.

In a CBER analysis of the SAS data set, there was balance between treatment arms, with all vials accounted for in 61% and 67% of subjects in the interferon β-1b and control groups, respectively. More than 80% of vials were accounted for in 77% and 82% of subjects in the interferon β-1b and control groups, respectively.

CBER also assessed accountability across study sites as a surrogate of general compliance with the protocol, and a pattern emerged. Based on an analysis of average per subject accountability ranked by site, accountability was clearly better than average at sites within the UK, and worse than average at German sites. Specifically, 9 of the 13 lowest ranked sites were in Germany. This is noteworthy not only because the UK and Germany are at the extremes of this compliance analysis, but because they are the two leading countries in terms of subjects enrolled.

Blinding

Blinding questionnaires were completed by Treating Physicians, EDSS Physicians and patients at the Month 36 visit or at premature termination. Respondents were asked to guess treatment assignment; however, a "don't know" selection was available to discourage random guessing. Data are available from ≈80% of subjects. Responses of "don't know" were approximately equally divided between subjects whose actual assignments were interferon and placebo; this answer was chosen by 67% of EDSS Physicians, 36% of Treating Physicians and 23% of subjects. The sponsor notes that blinding was most important for EDSS Physicians because they were responsible for assessing the primary outcome variable, and the clinical study report concludes that EDSS Physicians were well blinded because they could guess treatment assignment correctly only 19-22% of the time.

Reviewer's Comment:
A more conservative analysis considers that there is potential bias favoring a response of "don't know," because a correct guess suggests unblinding and potential bias on the part of study participants. If "don't know" responses are eliminated from the analysis, it can be seen that EDSS Physicians, treatment physicians and subjects guessed correctly 54-64%, 80-84% and 70-85% of the time, respectively, suggesting substantial unblinding for subjects and Treating Physicians, but reasonable blinding for the critical EDSS Physicians (Table 5).

Table 5: CBER Analysis of Blinding Questionnaire - "Don't Know" Responses Eliminated

guessed assignment		actual assignment	
		Placebo	Interferon
EDSS physician	Placebo	54%	36%
	Interferon	46%	64%
Treatment physician	Placebo	80%	16%
	Interferon	20%	84%
Subject	Placebo	70%	15%
	Interferon	30%	85%

Study Population: Baseline Characteristics

General

Baseline demographic characteristics and MS disease status are summarized by treatment group in Table 6. In general, the characteristics are typical of the SPMS patient population. Compared with the patient populations of prior clinical trials of interferons in subjects with RRMS, these subjects were older, had a longer total duration of MS, and exhibited a higher median baseline EDSS, as expected.

Deficiencies in data collection

For calculation of the duration of SPMS, no date of onset was reported in 25 interferon and 20 placebo subjects. In 6 subjects (3 in each group), the date of onset of SPMS was essentially the same as the study entry date.

The change in EDSS during the 24 months prior to enrollment was not reported in approximately 40% of subjects overall, although the lack of documentation was balanced between the groups (141 and 144 subjects in the placebo and interferon arms, respectively). The ramifications of this were discussed in the "Protocol Deviations" section (page 26).

Table 6: Demographic characteristics and baseline disease status of study population

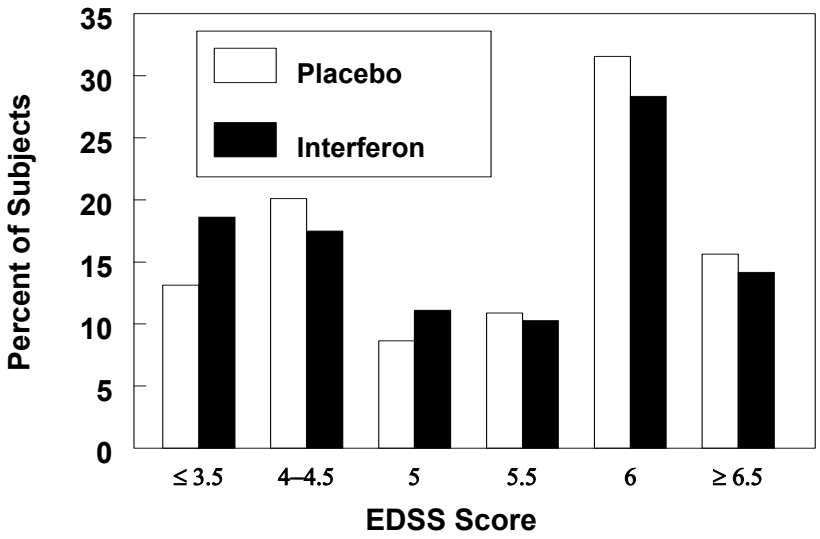
Characteristic	Interferon	Placebo
N	360	358
Age (years)		
median	42.0	41.7
range	22 - 57	21 - 56
% female	58.1	64.2
mass at start of study (kg)		
median	65.0	65.0
range	42 - 117	42 - 125
MS duration (years)		
median	11.7	11.9
range	1.0 - 36.3	1.9 - 40.2
Duration of SP phase (years)		
median	1.3	1.3
range	0 - 13.7	0 - 14.9
Numbers of patients with SP phase duration:		
< 1 year	74	81
1 year	126	122
2 years	58	52
3 years	34	39
≥ 4 years	68	64
Numbers of patients with EDSS score change over the 24 months prior to study:		
missing or not documented	144	141
<1	7	4
1.0	96	105
1.5	53	41
≥ 2	60	67
Baseline EDSS		
median	5.5	5.5
range	2.0 - 7.0	3.0 - 7.0
Numbers of patients with baseline EDSS:		
≤ 3.5	67	47
4.0 to 5.5	140	142
≥ 6.0	153	169
Numbers of patients by numbers of relapses in the 24 months prior to study:		
no relapses	112	97
1 relapse	53	56
≥ 2 relapses	195	205
MRI baseline T2 volume (cm^3)		
median	21.6	23.8
mean	26.5	28.0
SE	1.2	1.2
range	0.3 - 129	0.6 - 135

Balance between treatment groups

Generally, there was good balance of baseline characteristics between the treatment groups. Demographic characteristics were well-matched, with the exception of gender. There were more females in the placebo arm. MS tends to follow a more benign course in women; therefore, excess females in the placebo arm would be expected to bias the results against the active treatment, and this is not considered to be an important problem.

Balance was fair with respect to baseline disease status, with baseline status tending to be slightly worse in the placebo group. This conclusion is based on three key parameters:

Figure 1: Baseline EDSS by Treatment Group



a. Baseline EDSS - The pre-study EDSS score constitutes a key baseline parameter because it is indicative of the level of pre-existing disability, and progression of disability (the primary efficacy endpoint) is related to baseline disability. The distribution of baseline EDSS scores was not uniformly distributed across the scale range (Figure 1), but was similar to distributions in other MS studies. Epidemiological studies indicate that the duration at grade EDSS 6.0 is relatively long in comparison with grades 4.0 and 5.0, hence a greater frequency at EDSS = 6.0 is expected. With respect to balance between treatment groups, baseline EDSS tended to be slightly worse in the placebo arm in terms of the numbers of subjects with EDSS $\leq$ 3.5 (an excess of 20 subjects in the interferon group) and the numbers of subjects with EDSS $\geq$ 6.0 (an excess of 16 subjects in the placebo group, Table 6, Figure 1).

Reviewer's Comment:
In Figure 1, it is apparent that there are excesses of interferon subjects in two of the three baseline EDSS categories of 5.0 and below, whereas there are excesses of placebo subjects in the EDSS categories of 5.5 and above.

b. Relapses in the 24 months prior to study - In the placebo group, there tended to be fewer subjects with no relapses in the pre-study period, and more subjects with two or more relapses (Table 6).

c. MRI baseline T2 volume - The volume tended to be slightly greater in the placebo group (Table 6).

Individually, these differences in baseline EDSS distribution, relapse rate and T2 volume are minor; however, taken together, they are directionally similar and suggest the possibility of more severe baseline disease in the placebo group. If true, this would tend to bias the results in favor of the interferon group.

Study Results

Interim analysis

The interim efficacy analysis was planned to occur when all subjects completed 24 months on study. At the time of the interim efficacy analysis, placebo subjects had attained a mean time-on-study of 888 days, compared to 901 days for interferon subjects (overall mean = 895 days). For the final analyses, mean time-on-study was 1008 days for all subjects. Thus, the interim dataset included approximately 82% of the data that were planned to be available for the final analysis; the complete dataset included 92% of the planned data.

The sponsor considers the interim dataset to be the primary dataset for efficacy. Thus, data and analyses from the interim analyses were presented in extensive detail, whereas much detail is lacking with respect to the complete dataset analyses. (The sponsor considers the complete dataset the primary dataset for safety.)

The sponsor briefly summarized efficacy results from the complete dataset, and compared the statistical significance (interim versus complete data) on a number of endpoints. There was little difference with respect to the statistical analyses for the primary and secondary endpoints. The majority of the tertiary endpoint analyses were qualitatively unchanged, although four tertiary endpoints that were statistically significant in the interim analysis were non-significant in the final analysis. Conversely, two tertiary analyses changed from non-significant to significant. These changes were observed sporadically throughout several types of endpoints and in only a minority of the tertiary analyses, and were not considered to be important by the sponsor.

CBER deems it appropriate to include all available efficacy and safety data in its review of this sBLA. Thus, the results of interim analyses were not examined further and are not addressed in this review. The majority of analyses on the final dataset were performed by CBER. They are described in Times font.

Primary efficacy endpoint

In total, there were 356 progressions in the 718 subjects; 193 in the placebo group and 163 in the interferon group. Data were analyzed using Kaplan-Meier survival curves and tested using log-rank test. The sponsor reported a significant delay in time to progression for interferon-treated subjects. The extended Mantel-Haenszel test for covariance adjusted log-rank scores with stratification adjustment for center and covariance adjustment for center and baseline EDSS yielded a p-value of 0.0046. The p-value for the unstratified log-rank test was 0.0039. The sponsor summarized the results as time to disease progression at the 40th percentile (549 days for placebo, 901 days for interferon, p=0.0007).

Reviewer's Comment:
The use of this specific percentile was not prospectively specified in the statistical analysis plan. Use of the 40th percentile is unusual as a statistical approach, and this measure has limited clinical utility as a balanced representation of patient experience in this study.

CBER analyses

For the primary efficacy analysis on the complete dataset, the sponsor provided raw EDSS data with a ---------- program to output the results of the analyses in tabular form. The critical data for the Kaplan-Meier analysis (time to event and censoring status for each subject) could be obtained only through an intermediate ------- dataset which was produced by executing the ---------- program provided. CBER

performed an internal analysis of raw EDSS data taken directly from the original ------ transport dataset, verified against the case report forms. Time to progression and censoring were determined as directed in the protocol.

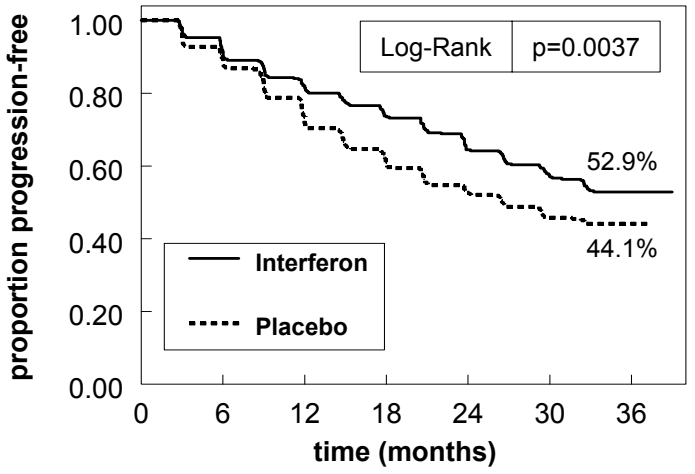
There was one notable difference between the CBER analyses and the sponsor's analyses. In the sponsor's analyses, the timing of events was categorized with respect to 3-month time periods, such that progression could only occur at Month 3, 6, 9, etc. In the CBER analyses, progression was analyzed by actual date, and Kaplan-Meier analyses were conducted using days as the unit of time. For comparability with the sponsor's analyses, CBER converted days to months using the a conversion factor of (365.25)/12

When the analyses of CBER and the sponsor were compared, there was only one "disagreement" with respect to time-to-progression. This occurred in a subject with a transient increase followed by a sustained increase in EDSS, and resulted in a change of time to progression of 161 days - a disparity that is clearly unimportant.

The results of the CBER analysis are shown in Figure 2. Using no adjustments for center or baseline EDSS, the p-value is 0.0037.

Based on the progression rates at 36 months, the Kaplan Meier analysis estimates annualized progression rates of 15.7% and 18.6% in interferon- and placebo-treated patients, respectively. **Thus, on average, interferon use was associated with an absolute 3% decrease in the annual rate of progression.**

Figure 2: CBER analysis of primary endpoint: time to disease progression



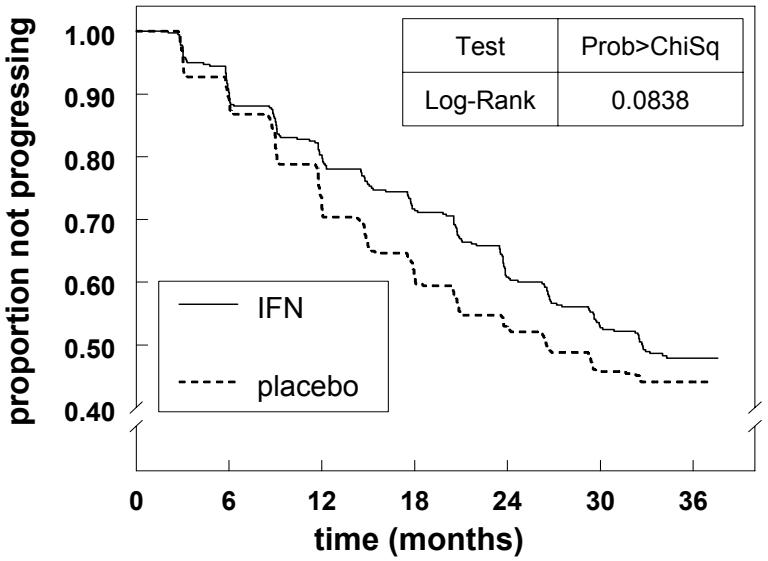
A non-parametric regression model (Cox model, proportional hazards fit) was used to test the hypothesis that demographic or baseline characteristics were significantly related to outcome. Treatment assignment was significantly related to outcome (risk ratio 0.86, 95% confidence limits 0.76 and 0.97). Age, gender, baseline EDSS, MS duration and SPMS duration were not significantly associated with outcome. Although baseline T2 volume was significantly associated with progression, the risk ratio was essentially unity, suggesting that this parameter was not clinically important.

Exploratory analyses on the primary endpoint

1. Subjects lost to follow-up

For the entire study, 88 subjects were lost to follow-up: 40 in the interferon group and 48 in the control group. For the primary endpoint, all subjects lost to follow-up were censored at the point of loss, that is, they were treated as though they did not progress. A "worst case" analysis was performed by CBER, which assumed that subjects in the interferon arm who were lost to follow-up actually met criteria for confirmed progression on the day they were lost, whereas control subjects who were lost to follow-up did not progress (Figure 3).

Figure 3: Time to Progression - CBER Analysis of "Worst Case" Handling of Subjects Lost to Follow-Up



This highly biased analysis adds 40 subjects to the 163 subjects who exhibited confirmed progression in the interferon group. The vast majority of these subjects did not have an unconfirmed EDSS increase at their last recorded visit; thus, they could not have progressed at the actual date they were lost to follow-up. This "worst case" analysis is supportive of a treatment effect, and shows that censoring of subjects lost to follow-up was unlikely to importantly affect the study results.

2. Arbitrary nature of a one-point increase in EDSS score

The EDSS scale is not linear with respect to impairment, and the clinical importance of lost abilities may not be consistent between adjacent 1.0 point increments. The requirement for a 1 point increase in EDSS as a definition of progression (a 0.5 point increase if baseline EDSS was 6 or greater) is arbitrary. The analysis for the primary endpoint was repeated using EDSS increases of 0.5 and 1.5 as definitions of progression. In both cases, 3-month confirmation of progression was required as per the original primary endpoint (Figure 4, Figure 5).

Figure 4: Time to Progression; Progression Defined as 0.5 Point Increase in EDSS

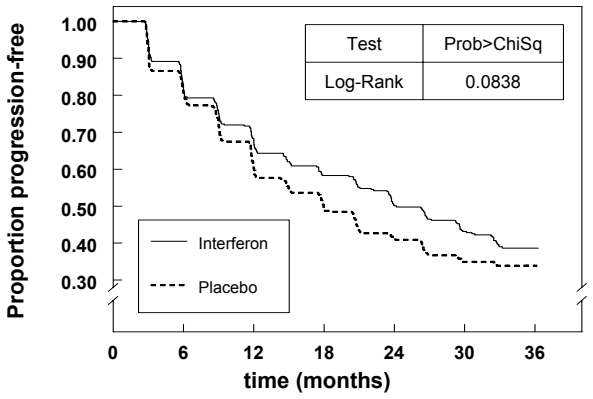
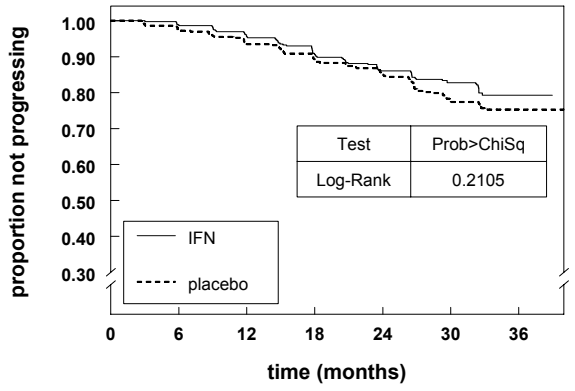


Figure 5: Time to Progression; Progression Defined as 1.5 Point Increase in EDSS



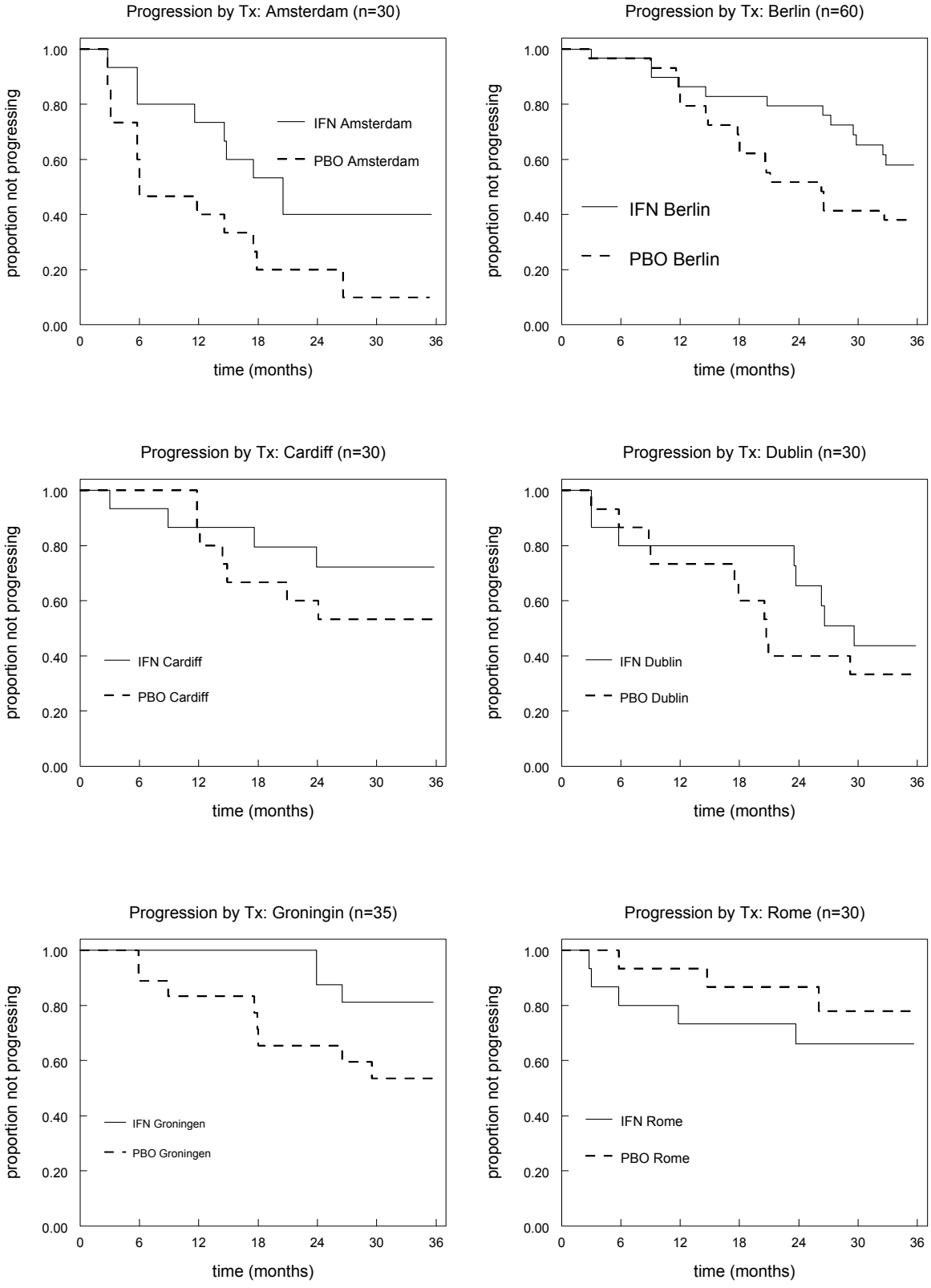
Using an EDSS increase of 0.5 units to define progression, events occur more commonly, as expected. The difference between groups does not reach statistical significance by log-rank; however, the trend is consistent with the results of the prospectively planned analysis using a one-point increase in EDSS.

With progression re-defined as a 1.5 point increase in EDSS, only 20-25% had a progression event, and the difference between treatment groups is not significant. It is of interest, therefore, that alteration of the definition of progression by the smallest possible unit in either direction would change the statistical outcome of the study.

3. Analysis of primary endpoint at major centers

There were six sites contributing 30 or more subjects to the study. Kaplan-Meier plots of time to progression were examined for these sites individually (Figure 6). Each plot represents a relatively small number of subjects; however, there was a trend towards a slower rate of progression in the interferon group in 5 of 6 of the largest centers. The trend was reversed at the Rome site, at which there were 5/15 progressions in the interferon group, and 3/15 progressions in the placebo group. Overall, this analysis demonstrates consistency in efficacy of interferon across the six largest contributing centers.

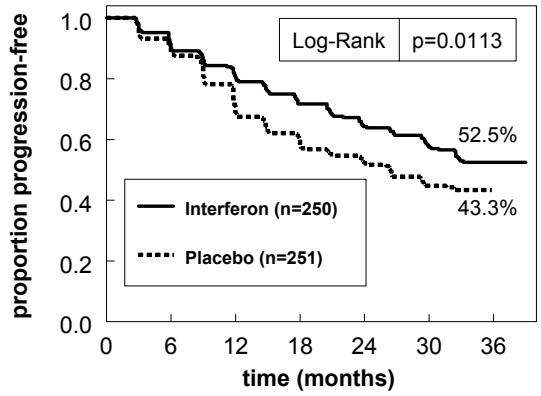
Figure 6: Analysis of primary Endpoint at Major Centers



4. Reliability of diagnosis and classification of SPMS

The reliability of the diagnosis and classification of SPMS is critical in assessing whether a new treatment is effective in this sub-population of MS patients. For this study, subjects were to have progressive deterioration as judged by the investigating physician, which was sustained for ≥ 6 months. A more strict definition of progressive deterioration would include a confirmed increase in EDSS in the pre-study period. In this study, 60% of subjects had documented increases in EDSS prior to enrollment, evenly divided between treatment groups. CBER performed a Kaplan-Meier time-to-event analysis on these subjects, and a significant treatment effect was demonstrable (Figure 7).

Figure 7: Time to progression: Subjects with documented increases in EDSS pre-study

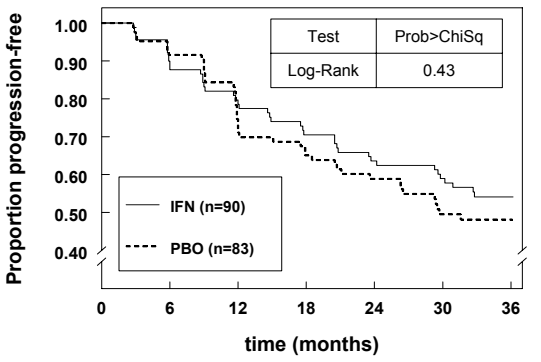


5. Potential effect of exacerbations on apparent progression

The efficacy of Betaseron has been established in RRMS. In order to expand the labeling to encompass patients with SPMS, the data should show that benefit is not due simply to a decreased frequency of relapses with diminished accumulation of relapse-related residual deficits.

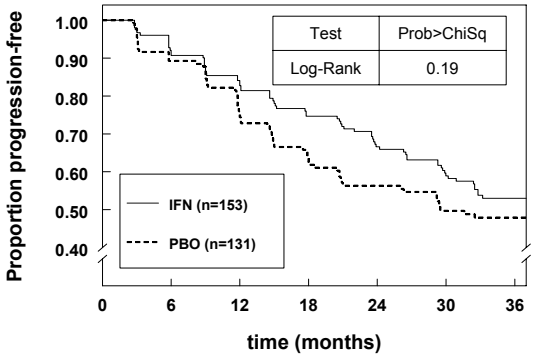
In an exploratory analysis (Figure 8), CBER assessed time to progression in the subset of subjects with documented progression but no reported exacerbations in the 24-month pre-study period. Although the number of subjects in this subset is limited, the results are directionally in favor of the active treatment.

Figure 8: Subjects with no exacerbations in the pre-study period



An additional CBER time to event analysis in subjects who had no exacerbations during time-on-study provides an additional means of addressing this issue (Figure 9). Subjects in this analysis *who did progress* were unlikely to have done so as a result of accumulation of relapse-related deficits. As shown in Figure 9, there is a trend in favor of interferon in subjects with no on-study exacerbations, supporting the concept of a treatment-associated delay in progression that is independent of exacerbations.

Figure 9: Subjects with no exacerbations during time-on-study



6. Use of concomitant medications

Numerous medications have been used for symptomatic benefit in MS. Although the use of these agents has not been shown to affect the natural course of the disease, concomitant medication use has to potential to alter symptoms, perception of symptoms, and/or physical findings in MS. Thus,

concomitant medication use, if unbalanced in the study, could confound the results by differentially affecting the EDSS. Agents with the potential to importantly influence patient assessments are summarized below:

- Limb spasticity, ataxia and weakness are amenable to pharmacologic modification with muscle relaxants, antispasmodics, and benzodiazepines, and such modification could be sufficiently effective to alter the EDSS score of a subject.

- Fatigue is a frequent symptom which may alter mobility in MS patients, thereby affecting EDSS. Dopaminergic agents (typically amantadine) and psychostimulants (typically pemoline) have been used to reduce fatigue in this disease.

- Bladder and bowel function are components of Kurtzke Functional Systems (FS) assessments, which are, in turn, components of EDSS assessments. Therefore, use of parasympathomimetic and antiadrenergic agents are potentially confounding variables.

- Numerous agents prescribed for neuropsychologic impairment and pain management, including benzodiazepines, antidepressants and anticonvulsants, also can affect FS assessments and EDSS.

- Prohibited immune modifiers such as glatiramer acetate, interferon β-1a and interferon β-1b could confound the results of the study if their frequency of use was significant and unbalanced.

The CBER analyses of concomitant medication use were based on the 50,117 line listings of concomitant medications from the final data set. Data from all subjects was included and subjects were coded as randomized (Table 7). There was greater use of immunomodulatory agents and NSAIDs in the interferon group. These differences will be discussed below. There was increased use of steroids in the placebo group. If steroids are beneficial in MS, then excess use in the placebo group would bias the results against interferon.

Table 7: Concomitant Medications - Months of Use Per Subject; Interferon vs. Placebo (Mean ± SEM)

Agent	Interferon	Placebo
N	360	358
NSAIDs	10.9 ± 0.7	6.4 ± 0.6
muscle relaxants	10 ± 0.8	9.6 ± 0.7
antidepressants	6.6 ± 0.6	6.1 ± 0.6
benzodiazepines	5.3 ± 0.6	5.4 ± 0.6
steroids	2.4 ± 0.3	3.2 ± 0.3
antiepileptics	2.2 ± 0.4	3.3 ± 0.5
opiods	1.3 ± 0.3	1.4 ± 0.3
psychostimulants	1.6 ± 0.3	1.5 ± 0.3
parasympathomimetics	1.1 ± 0.2	1.0 ± 0.1
immunomodulators	0.5 ± 0.2	1.2 ± 0.2

Differences in concomitant medication use were further analyzed by evaluating use between progressors and non-progressors in the two treatment arms (Table 8). Although imbalanced between treatment arms, NSAID use was, on average, similar in progressors and non-progressors in both treatment arms. (The potential effects of NSAIDs on progression are analyzed further, below.)

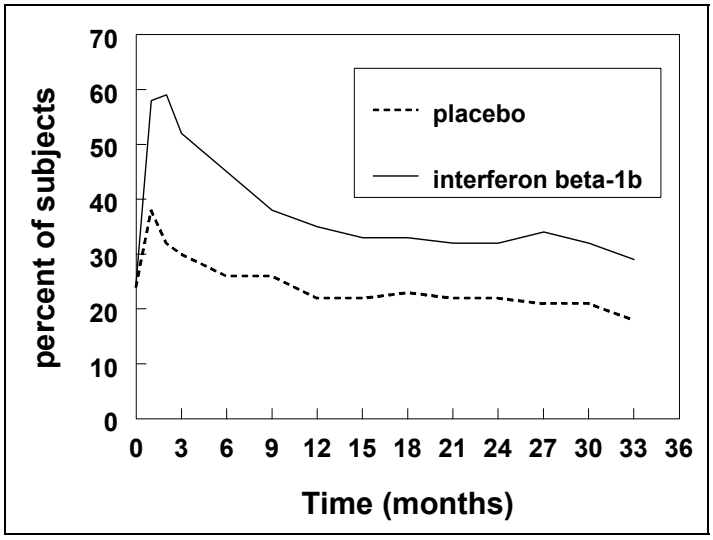
Table 8: Concomitant Medications - Months of Use Per Subject; Progressors vs. Non-Progressors; Interferon vs. Placebo (Mean ± SEM)

	Interferon		Placebo	
Agent	non-Progressors	Progressors	non-Progressors	Progressors
N	197	163	165	193
NSAIDs	10.8 ± 1	11 ± 1	6.2 ± 0.8	6.6 ± 0.8
muscle relaxants	7.5 ± 0.9	13 ± 1.2	8.4 ± 1.1	10.6 ± 1
antidepressants	5.4 ± 0.8	8 ± 1	5.5 ± 0.9	6.6 ± 0.8
benzodiazepines	5.2 ± 0.8	5.6 ± 0.9	5.5 ± 0.9	5.4 ± 0.8
steroids	1.9 ± 0.3	3.1 ± 0.5	2.8 ± 0.4	3.5 ± 0.4
antiepileptics	2.3 ± 0.5	2 ± 0.5	2.2 ± 0.6	4.2 ± 0.7
opiods	1.3 ± 0.4	1.2 ± 0.4	1.7 ± 0.6	1.1 ± 0.4
psychostimulants	1.5 ± 0.4	1.8 ± 0.5	1.5 ± 0.4	1.5 ± 0.4
parasympathomimetics	1.2 ± 0.2	1.0 ± 0.2	1.0 ± 0.1	1.1 ± 0.2
immunomodulators	0.2 ± 0.1	1 ± 0.3	0.4 ± 0.2	1.9 ± 0.4

Conversely, excess use of immunomodulatory agents was associated with progression within each treatment arm. Most of this excess use was attributable to methotrexate, with reported use in 30 subjects in the placebo group and 3 subjects in the interferon group. If methotrexate has salutary activity in MS, the excess use in the placebo arm would tend to bias the results against interferon. There is excess steroid use in placebo subjects as noted above, with excess use in progressors in both arms. There was also excess use of anti-epileptic agents in progressors in the placebo arm.

In a CBER analysis, NSAID use was further characterized by plotting use by study month (Figure 10). There was excess NSAID use in the interferon arm by approximately a 3:2 ratio throughout the study. Though all subjects were encouraged to use ibuprofen initially to reduce fever-like symptoms and improve blinding, at most, only 59% of subjects used NSAIDs in the interferon arm, and only 38% of subjects used NSAIDs in the placebo arm. It is plausible that the decreases in NSAID use between Month 2 and Month 12 were related to a perceived lack of need or lack of NSAID effectiveness on the part of subjects.

Figure 10: NSAID Use by Treatment and Time

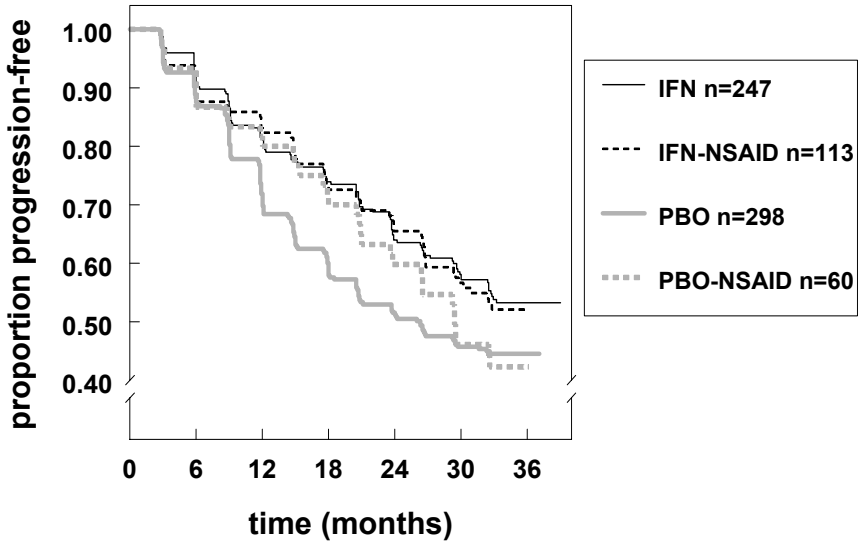


In light of the substantially greater use of NSAIDs in the interferon arm, time to progression was assessed in subgroups of NSAID use and treatment assignment in an exploratory analysis (Figure 11). For this analysis, the criterion for NSAID use was the listing of any NSAID for at least six of the 3-month reporting intervals in the concomitant medications dataset. This corresponds to approximately 18 months

or more of use. This definition of NSAID use, although arbitrary, placed 31% of interferon subjects and 17% of placebo subjects in the NSAID+ category, which appears to be consistent with the data as shown in Figure 10.

Although the number of subjects in the placebo arm with NSAID use is limited, NSAID use appears to be negatively associated with progression within the placebo arm. (Note that this post-hoc exploratory analysis can only suggest an association and can not address causality.) Nevertheless, in light of the modest imbalance in NSAID use and the modest strength of association, it is unlikely that NSAID use was a major confounding factor in this study.

Figure 11: Time to Progression: Effect of NSAID Use

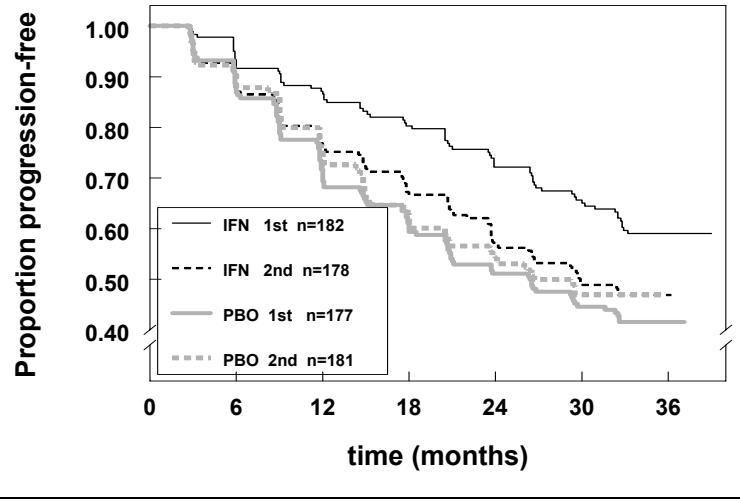


An additional analysis of the potential effects of these agents on disability progression was made by determining the proportions of subjects progressing within each treatment arm and concomitant medication category. Overall, no specific patterns of imbalance in concomitant medication use emerged which were thought to be capable of importantly confounding interpretation of the primary efficacy endpoint.

7. Subgroups defined by time of entry into study

CBER performed an exploratory analysis on the primary efficacy endpoint with respect to time of study entry. Subjects were divided into two groups based on study entry before or after the median enrollment date ("1st" vs. "2nd," Figure 12). There is a robust treatment effect in the 359 subjects enrolled in the "first half" of the study; however, there is no apparent treatment effect in the second half of the study (359 subjects).

Figure 12: Time to progression: Subgroups defined by time of study entry



Such a disparity could be a manifestation of temporally-related changes in the therapy of MS (i.e., the emergence of a new treatment that alters the natural history of the

disease); however, this is unlikely to be operative in this study because enrollment commenced at all sites within an 11-month period.

Differences in baseline demographic characteristics and/or disease activity provide other potential explanations for this observation. CBER analyzed baseline demographic and disease-specific variables by time of study entry (first/second half) and treatment assignment. Analyzed using the Wilcoxon test, baseline EDSS score in the "first-half" interferon sub-group, the sub-group for which time-to-progression was delayed, was significantly lower than the other sub-groups (p=0.01, Table 9). This finding appeared to be related to excess subjects within the lowest baseline EDSS category (EDSS ≤ 3.5, Table 9). This imbalance may be consequential, because it indicates an excess of subjects with less advanced disease within the subgroup of subjects that "drove" the efficacy endpoint. The "first-half" interferon subgroup was also notable for an excess of male subjects and shorter baseline MS duration relative to the other three sub-groups, although these differences were not statistically significant (Table 10). Given that these were only modest imbalances with respect to factors that are not strongly associated with progression, they were unlikely to importantly affect the results in this subgroup.

Table 9: Baseline EDSS by treatment assignment and time of study entry (CBER analysis)

	Baseline EDSS	First half of enrollment		Second half of enrollment	
		Placebo	Interferon	Placebo	Interferon
numbers of subjects:	≤ 3.5	28	46	19	21
	4 - 4.5	35	31	37	32
	5 - 5.5	38	35	32	42
	6	49	48	64	54
	> 6	27	22	29	29
baseline EDSS:	median	5.5	5.0	6.0	5.5
	mean	5.1	4.9	5.3	5.3

It is important to note that enrollment did not proceed simultaneously at all sites. Some sites tended to enroll subjects at the beginning of the study, others towards the end. Therefore, the apparent effect of time of enrollment on outcome could be related not to temporal factors per se, but to regional factors or differences between centers. Such differences could be subtle, and could include disparate patient characteristics, practice of medicine, concomitant medication use, patient assessments, and data interpretation. Thus, CBER analyzed time of enrollment and the primary efficacy endpoint, time to progression, by country.

Table 10: Baseline gender and MS duration by treatment assignment and time of study entry (CBER analysis)

	First half of enrollment		Second half of enrollment	
	Placebo	Interferon	Placebo	Interferon
% male	34.5%	45.6%	37.0%	38.2%
median MS duration (years)	11.8	11.0	12.4	12.2

8. Subgroups defined by country

To assess the possibility of differential efficacy by country, CBER assessed time of enrollment by center and country. Because of the relatively large number of sites, the analysis by country is more informative. Table 11 shows patient enrollment by country and time (first versus second half), as well as *excess* progression events in the placebo group, also by half of study. Belgian, Swiss, German and Spanish centers enrolled subjects predominantly in the first half of the study; Italian, British and Finnish sites

generally enrolled subjects in the second half. Efficacy in the first half of the study was driven primarily by excess progressions in the placebo group at sites in France, UK and the Netherlands. These three countries together accounted for 20 of the 29 excess progressions in the placebo group. Consistent with Figure 12, the total progressions in the interferon and control arms during the second half of enrollment were essentially equal.

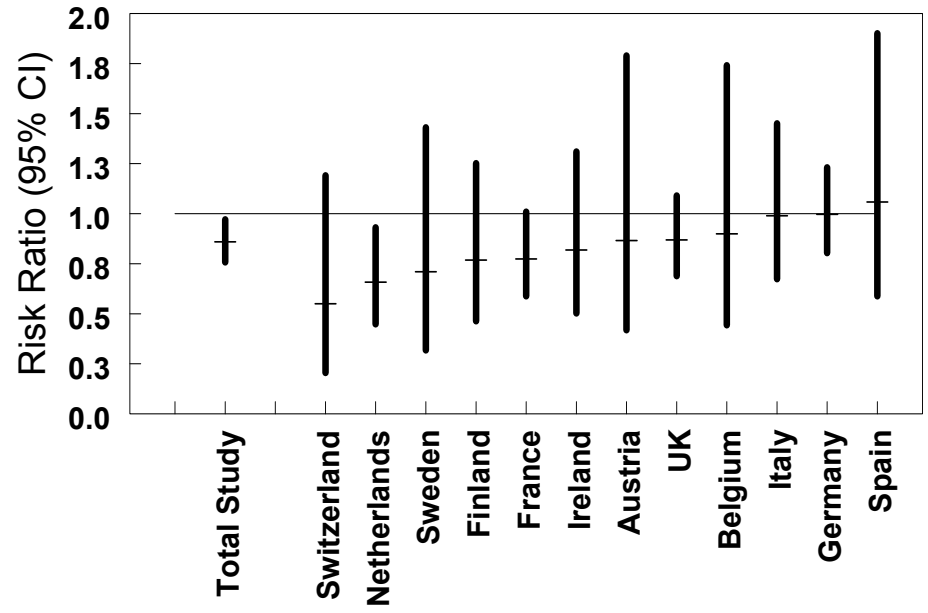
Country-specific risk ratios with 95% CI for the primary time-to-progression efficacy endpoint are shown

Table 11: Patient enrollment, excess progression events by country and time

Country	Patients Enrolled		Excess Progression Events in Placebo Group	
	First Half	Second Half	First Half	Second Half
Austria	4	8	-1	1
Belgium	24	0	1	0
Finland	0	36	0	3
France	57	48	10	-2
Germany	126	38	0	-4
Ireland	18	12	2	0
Italy	18	54	3	-2
Netherlands	34	31	5	4
Spain	17	7	1	-1
Sweden	7	11	0	2
Switzerland	18	0	3	0
UK	36	114	5	0
Total	359	359	29	1

in Figure 13. German sites accounted for the greatest number of subjects in the study, and enrolled predominantly in the first half of the study. Interestingly, there was no interferon treatment effect apparent in that country (risk ratio ≈ 1). It can not be determined whether the lack of efficacy on the

Figure 13: Risk Ratios and 95% Confidence Intervals for Time to Progression by Country



progression endpoint at German sites is a chance finding, or due to important differences in patient characteristics, management strategies, or assessment.

9. Masking of EDSS scores during exacerbations for the time-to-progression analyses

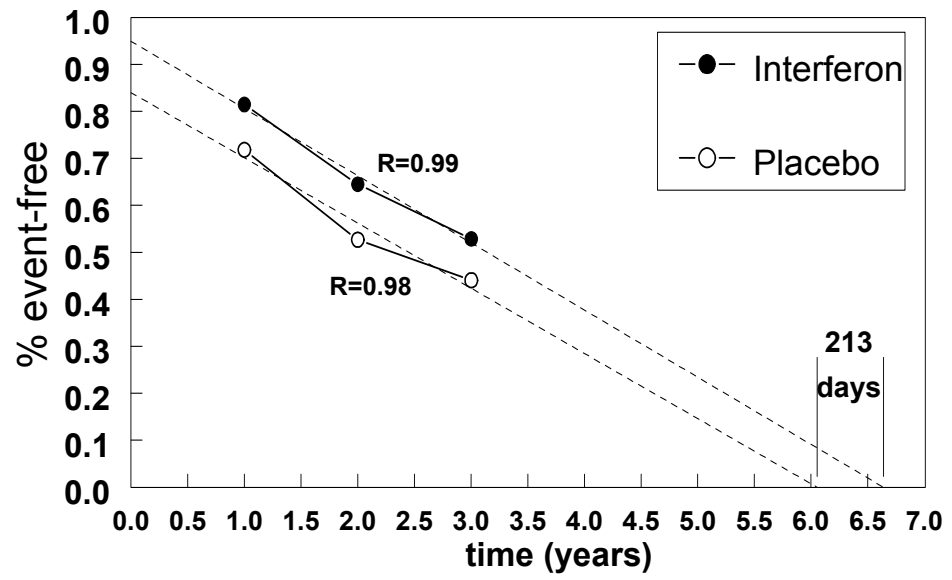
Masking of EDSS scores during relapses for the time-to-progression analysis has the potential to delay or eliminate progressions. An excess of relapses in the placebo arm, as would be expected on the basis of the known effects of interferon, would tend to delay the apparent rate of EDSS progression in that treatment group. CBER conducted an exploratory analysis to address this issue.

There were 233 and 287 blanked EDSS scores in the interferon and placebo groups, respectively. The elimination of blanking shortened the time to progression in 18 subjects in the active treatment arm, and 33 subjects in the placebo arm. Four subjects in each treatment group who did not progress in the original analysis would be re-classified as progressors. Thus, as expected, the blanking of EDSS scores during exacerbations did not impart an advantage to interferon in this analysis. Conversely, the apparent treatment effect of interferon would have been augmented if EDSS scores had not been blanked during exacerbations.

10. Duration of treatment effect

In an exploratory analysis, CBER examined more closely the Kaplan-Meier estimates of progression rates at 1, 2, and 3 years (Figure 14). For both treatment groups, we found a linear relation between event-free survival and time from Year 1 through Year 3 (using least-squares, the R-values are 0.99 and 0.98 for the interferon and placebo groups, respectively). Interestingly, for the two treatment groups, the slopes of the lines, and therefore the event rates, are essentially the same from Year 1 through Year 3. This suggests that the benefit of interferon is realized only during the first year of treatment. Because the risk of interferon administration is not inconsequential, this raises the question of the optimal duration of treatment, an issue that might be explored in a phase 4 study.

Figure 14: Linear relation between event-free survival and time from Year 1 to Year 3



Secondary Endpoints

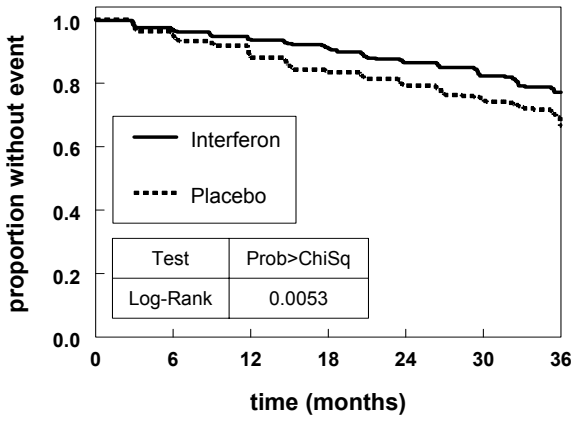
1. Time to becoming wheelchair-bound

For the interim analysis reported by the sponsor, there were 88 and 59 events in the placebo and interferon groups, respectively. Using a Mantel-Cox log-rank test with stratification adjustment for time interval and baseline EDSS, the reported p-value was 0.0142 for time to wheelchair-bound. The sponsor noted potentially significant treatment by gender and body surface area interactions which suggested a greater treatment effect for females and subjects with body surface area <1.68 m^2. For males, there were 23 events in 128 subjects (18.0% with event) in the placebo group, whereas there were 30 events in 151 subjects in the interferon group (19.9% with event). The sponsor questioned the significance of these findings, however, because they found no differential treatment effect with respect to the primary endpoint.

Exploratory analyses on time to becoming wheelchair-bound endpoint

CBER conducted exploratory analyses on the complete dataset. For the complete dataset, there were 109 and 77 events in the placebo and interferon groups, respectively. The Kaplan Meier time-to-event analysis is shown in Figure 15. The non-adjusted p-value is 0.0053. The requirement for a wheelchair is less subjective than criteria for EDSS progression, and additional exploratory analyses were not conducted. Analyses of subgroups for differential effects is presented later in this review.

Figure 15: Progression to Wheelchair-Bound: CBER Analysis, Complete Dataset



2. Annual relapse rate

The sponsor reported relapse rates as of the interim analysis (Table 12). The p-value for the Extended Mantel-Haenszel test with covariance adjustment for center was 0.0006. The treatment effect was also highly significant when comparing subjects either with or without relapses in the 24 month pre-study period.

Table 12: Annual Relapse Rates by Exacerbation Severity - Interim Data

Relapse Severity	Placebo mean ± SD	Placebo median	Interferon mean ± SD	Interferon median
moderate or severe	0.49 ± 0.79	0.33	0.31 ± 0.49	0.00
all	0.63 ± 0.88	0.39	0.42 ± 0.59	0.36

Exploratory analyses on the relapse endpoint

Analyses of the complete data set were performed by CBER using raw data in ----- transport datasets. A total of 1077 investigator-verified relapses were reported, with 464 in interferon-treated subjects and 613 in controls. There were also 140 relapses lacking a valid start date, 77 were in the interferon group; 63 were in the placebo group. These relapses were considered unconfirmed. The annual relapse rate was calculated as the number of relapses divided by time-on-study. Annual exacerbation rates were assessed by exacerbation severity (Table 13).

Table 13: CBER Analysis of Annual Relapse Rates by Exacerbation Severity - Complete Dataset, Confirmed Exacerbations Only

Relapse Severity	Placebo		Interferon	
	mean ± SD	median	mean ± SD	median
mild	0.15 ± 0.36	0.00	0.12 ± 0.26	0.00
moderate	0.41 ± 0.72	0.00	0.28 ± 0.52	0.00
severe	0.14 ± 0.44	0.00	0.09 ± 0.34	0.00
moderate or severe	0.55 ± 0.98	0.34	0.37 ± 0.65	0.00
all	0.7 ± 1.05	0.34	0.49 ± 0.73	0.34

The effect of Betaseron on annual relapse rate is directionally similar in all three severity categories, with decreases in mean annual exacerbation rate in the 20-40% range. CBER also assessed annual relapse rates with the inclusion of unconfirmed exacerbations. Exacerbation rates were modestly higher, and the results were directionally similar (Table 14).

Table 14: CBER Analysis of Annual Relapse Rates by Exacerbation Severity - Complete Dataset, Confirmed and Unconfirmed Exacerbations

Relapse Severity	Placebo		Interferon	
	mean ± SD	median	mean ± SD	median
mild	0.17 ± 0.37	0.00	0.14 ± 0.27	0.00
moderate	0.43 ± 0.72	0.34	0.32 ± 0.56	0.00
severe	0.15 ± 0.44	0.00	0.1 ± 0.35	0.00
moderate or severe	0.59 ± 0.99	0.34	0.43 ± 0.7	0.33
all	0.76 ± 1.05	0.37	0.57 ± 0.78	0.34

An analysis of worst relapse per subject showed that for the interferon group, there was a net shift of approximately 20 subjects from the "severe" category to the "none" category (Table 15).

Table 15: Severity of Worst Relapse

	Placebo	Interferon
none	131	154
mild	33	36
moderate	122	117
severe	72	53
total	358	360

3. MRI - Percentage change in T2

lesion volume

T2 lesions are thought to represent fixed disease due to prior attacks. The T2 volume represents an approximation of the total extent of these lesions, which is thought to reflect the overall burden of disease. Appropriately, an assessment of the correlation between lesion volume and activity and clinical benefit was a phase 4 commitment of the sponsor. MRI scans were scheduled at baseline and at the end of each year on study, to be performed within ±20 days of scheduled visits. Scans were not obtained during a clinical relapse or while subjects were receiving systemic corticosteroids. All scans were evaluated by a central evaluating center. The sponsor's analysis of the interim dataset is summarized in Table 16. The pre-specified analysis for efficacy was change from baseline to last scan using an extended Mantel-Haenszel test with stratification adjustment for center. The p-values reported were obtained using covariance adjustment for baseline lesion volume, as well as stratification adjustment for center. Month 36 data are quite limited as of this interim analysis. Whereas there was, on average, a relative increase in lesion volume in placebo-treated subjects, there was no apparent increase in mean lesion volume in Betaseron-treated subjects.

Table 16: MRI T2 Volume by Year - Sponsor's Analysis of Interim Data

	Baseline*		Month 12		Month 24		Month 36		Last Scan	
	Placebo	Interferon	Placebo	Interferon	Placebo	Interferon	Placebo	Interferon	Placebo	Interferon
***N* with scans**	344	346	320	329	301	306	28	25	327	333
Percent Change (%)										
mean			3.62	-3.71	7.83	-4.67	8.53	1.31	7.82	-4.88
median			1.66	-4.94	2.52	-6.90	6.35	-11.38	2.96	-6.91
P-value			<0.0001		<0.0001		0.042		<0.0001	
Absolute Change From Baseline (cm^3)										
mean	28.4	26.6	1.31	-1.22	2.32	-1.49	1.31	-1.88	2.27	-1.47
median	23.8	21.6	0.32	-0.77	0.42	-1.05	0.67	-2.12	0.46	-1.06
P-value			<0.0001		<0.0001		0.052		<0.0001	

* baseline volume

Exploratory analyses on the T2 volume endpoint

CBER reviewed the complete MRI datasets. MRI scans were obtained successfully for 87% of the planned evaluations, and MR data were available for >75% of subjects at Month 36. Baseline T2 volume was well balanced between the treatment arms (Figure 16).

The CBER analysis of the final dataset is in substantial agreement with the sponsor's interim analysis (Table 17). The results were qualitatively similar when analyzed as percent change or as actual T2 volume as a function of time.

Figure 16: Baseline T2 Volume by Treatment Group

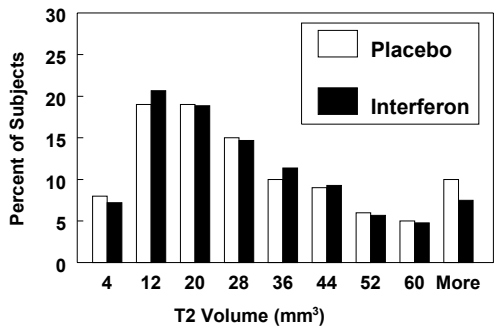


Table 17: MRI T2 Volume by Year - CBER Analysis of Final Data

	Baseline		Month 12		Month 24		Month 36	
	Placebo	Interferon	Placebo	Interferon	Placebo	Interferon	Placebo	Interferon
***N* scans**	335	332	321	329	302	308	274	293
Percent Change (%)								
mean			3.6	-3.7	7.5	-4.7	16.0	-1.6
median			1.6	-4.8	1.7	-6.9	11.0	-5.2
Absolute T2 Volume (mm3)								
mean	28.0	26.5	29.2	25.6	30.9	25.0	32.3	25.7
median	23.8	21.6	23.4	20.6	25.4	19.7	25.5	21.2

For both absolute and relative (percent change from baseline) data, there appeared to be a lack of accumulation of lesion volume in interferon-treated subjects over the 36-month study period, whereas an increase in T2 volume was observed in the placebo group.

4. Number of newly active MRI lesions, Months 1-6, Months 18-24

To assess the effect of interferon on new lesion formation, a 125-subject frequent MRI subgroup underwent exams before and after administration of a gadolinium contrast agent at monthly intervals from Month 1-6 and from Month 18-24. These were performed in addition to the annual exams. Month 0 constituted the baseline for the Month 1-6 examinations; Month 18 was the baseline for the Month 19-24 examinations. Included were lesions present on T2-weighted or gadolinium enhancing images which displayed new enhancement relative to baseline, or were non-enhancing but new or increased in volume on the T2 scan. Subjects with missing baseline scans were non included in the analysis. The sponsor included only the Month 1-6 data as a secondary endpoint. Data were reported on 60/61 subjects in the placebo group and 63/64 subjects in the interferon group (Table 18).

Table 18: Cumulative number of newly active MRI lesions in Months 1-6 - Sponsor's interim analysis (mean number ± SEM)

	Placebo	Interferon
Month 1	2.40 ± 0.56	0.92 ± 0.27
Month 2	4.07 ± 0.83	1.30 ± 0.42
Month 3	6.08 ± 1.17	1.86 ± 0.65
Month 4	7.53 ± 1.41	2.40 ± 0.90
Month 5	9.11 ± 1.68	2.87 ± 1.06
Month 6	10.6 ± 1.95	3.53 ± 1.30

The CBER analysis of the complete dataset is substantially in agreement with the sponsor's analysis (Table 19). Interferon suppressed or prevented the formation of newly active MRI lesions by a factor of approximately 3:1 during the first 6 months of treatment. Importantly, the data also demonstrate persistent treatment effects of interferon, with suppression of newly active lesion formation by a factor of 4-5:1 during months 19-24.

Table 19: Cumulative number of newly active MRI lesions: CBER analysis of complete dataset (mean number ± SEM); n = number scans at each time point

	Placebo	*n*	Interferon	*n*		Placebo	*n*	Interferon	*n*
Month 1	2.36 ± 0.55	61	0.92 ± 0.27	64	**Month 19**	1.74 ± 0.43	50	0.4 ± 0.24	53
Month 2	4.02 ± 0.82	61	1.31 ± 0.41	64	**Month 20**	2.8 ± 0.61	51	0.63 ± 0.42	49
Month 3	6.07 ± 1.17	60	1.86 ± 0.65	63	**Month 21**	3.98 ± 0.92	49	0.79 ± 0.46	52
Month 4	7.05 ± 1.35	59	2.34 ± 0.91	62	**Month 22**	4.56 ± 1.05	50	1.16 ± 0.64	50
Month 5	8.43 ± 1.62	58	2.87 ± 1.06	63	**Month 23**	5.74 ± 1.41	47	0.78 ± 0.41	50
Month 6	10.2 ± 1.9	58	3.45 ± 1.32	62	**Month 24**	7.12 ± 1.57	51	1.55 ± 0.69	56

Tertiary Endpoints

EDSS Endpoints

1. Proportion of subjects with confirmed disease progression

For the interim dataset, the sponsor reported progression rates of 39% and 50% in the interferon and placebo groups, respectively (p=0.0038)

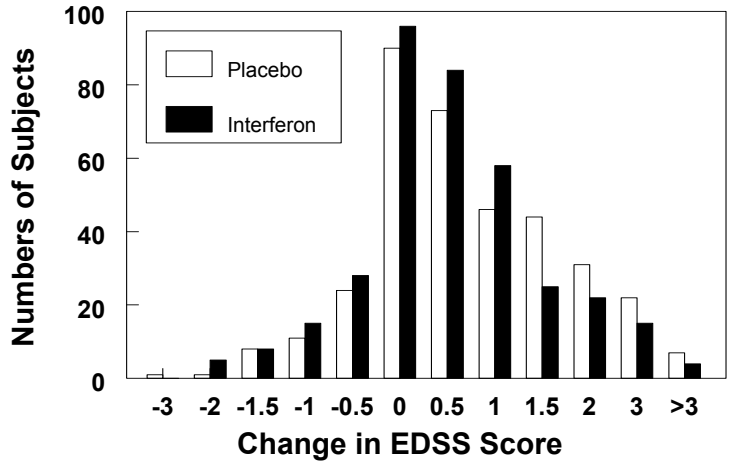
In a CBER analysis of the complete data set, there was confirmed EDSS progression in 45% of subjects in the interferon arm and 54% of subjects in the placebo arm (p=0.0251, Fisher's Exact Test; two-tailed). The higher rates of progression in the CBER analysis result from the accumulation of additional data after the interim cut off date.

2. Change in EDSS from baseline

For the interim data set, the sponsor provided an analysis of change in EDSS from baseline, where change was classified into five categories (<0.5, 0.5, 1, 1.5, ≥2). Using an extended Mantel-Haenszel statistic stratified for baseline EDSS, the difference between treatment arms was statistically significant (p=0.0278).

CBER performed a similar analysis on the final data set of unconfirmed EDSS scores. Using 0.5 EDSS point categories between -2 and 2, a significant treatment effect was demonstrable. A histogram of changes in EDSS for the two treatment arms is shown in Figure 17. A "leftward shift" (towards less progression) is apparent in the interferon treatment arm. Specifically, there tend to be greater numbers of subjects from the interferon arm in each delta EDSS category of -3 (improvement) to +1; there are more placebo-treated subjects in each delta EDSS category of 1.5 or greater.

Figure 17: Distribution in Change in EDSS Score from Baseline



3. EDSS at end-study

From the interim dataset, the sponsor analyzed the last available EDSS scores of the interferon and control groups, and found mean EDSS scores of 5.57 versus 5.84, respectively (p=NS).

CBER performed a similar analysis on the complete dataset. Final EDSS was analyzed using a CMH test stratified for baseline EDSS. In the CBER analysis, the effect of interferon was statistically significant (p=0.0224).

MRI endpoints

The sponsor performed numerous supportive analyses on the annual MRI data, and all showed a highly significant interferon treatment effect. These analyses included: absolute change in T2 volume, number of new or enlarged lesions, proportion of active scans/subject and the proportion of subjects with at least one active annual scan. Additional analyses for the frequent

MRI subgroup (months 1-6; months 19-24) included: number of persistently enhancing lesions, proportion of active scans and proportion of subjects with active scans. A significant treatment effect was demonstrable. For these analyses overall, there was some loss of power for months 19-24 compared to the first six month interval due to the loss of 16 subjects from the cohort. Nonetheless, the treatment effect in months 19-24 was highly significant. Of note, interferon treated subjects had lower accumulation of persistent lesions at every monthly time point in the first six months, consistent with a rapid onset of effect. The protocol also specified that enhancing lesion load in Months 1-6 and Months 19-24 would be tertiary endpoints; however, this analysis had not been performed at the time of sBLA submission and is pending.

Relapse endpoints

Based on the sponsor's analyses of the interim dataset, there were statistically significant treatment effects of interferon with respect to time to first relapse, the proportion of patients with one or more relapses, and relapse severity. There was no significant effect of interferon on the mean duration of individual relapses (placebo 55 days; interferon 52 days); however there was a significant treatment-associated reduction in mean total number of days spent in relapses throughout the study period (placebo 131 days, interferon 101 days; p=0.0096). There was also a reduction in number of days spent in moderate or severe relapses.

Cognitive function endpoints

A total of 476 subjects underwent cognitive function testing consisting of ten analyses performed for the Brief Battery of Neuropsychological Tests in MS. With the exception of one test, the effect of interferon was not statistically significant.

Quality of Life: Sickness Impact Profile Variables

The Sickness Impact Profile (SIP) Variables were used to assess sickness-related dysfunction in 12 areas of activity at 6-month intervals. Baseline scores were similar between the two groups. Results were not consistent for every time point for every test, and the difference in the overall score was not statistically significant. There was less worsening in the physical dimension score in interferon-treated subjects (body care and movement, ambulation, mobility, p=0.0305); however, the differences in the three individual scores were not individually significant. There was no significant treatment effect on the change in psychosocial dimension (emotional behavior, affective behavior, social interaction, and communication).

For the overall changes in the 12 individual scores, only the "eating" score was significantly better in interferon-treated subjects.

Ambulation Index

The Ambulation Index score is indicative of overall mobility from 0 (asymptomatic) to 9 (wheelchair-bound, unable to transfer independently). Interferon subjects had significantly lower AI scores at the last visit available for the interim analysis (interferon 4.24, placebo 4.69; p=0.032). Although more interferon subjects had minimal deterioration ($\leq$ 2 points), or remained stable or improved, the difference was not statistically significant. Analyses of time to $\geq$ 2-point worsening and proportion of subjects worsening by $\geq$ 2 points were also not significant. CBER did not perform analyses on the complete dataset.

GEMS

The sponsor used the complete dataset for the primary analysis, because very limited data were available at the time of the interim analysis. For this evaluation, disease severity was

subjectively assessed on a 7-point scale (very much better to very much worse) at baseline and at Month 36, with subtraction of the values to assess overall change in disease. Using extended Mantel-Haenszel statistics with stratification adjustment for center, changes in disease, neurological impairment, and disability were each statistically significant.

Steroid use

The sponsor's analysis of the interim dataset demonstrated significantly greater use of steroids in the placebo group, both were assessed as proportion of patients with steroid use (placebo 71%, interferon 57%; p<0.0001, Mantel-Haenszel test with stratification adjustment for center) or number of steroid courses per patient (placebo 1.91, interferon 1.34, p<0.0001, extended Mantel-Haenszel test with stratification adjustment for center).

The CBER analysis of the final dataset is summarized in Table 20. There was significantly greater steroid use in the placebo group when assessed as number of subjects with steroid use, steroid courses per subject and total steroid dose.

Table 20: CBER analysis of steroid use - final dataset

	Interferon	Placebo	p-value
N	224	261	0.0024*
% of total	62.8%	72.8%	
Steroid courses per subject			
mean	1.62	2.11	0.0011**
SEM	0.12	0.13	
Total steroid dose per subject (mg)			
mean	4825	6038	0.0012**
SEM	364	378	
Steroid protocol violations			
n	23	27	
% of total	6.4%	7.5%	

* Fisher's Exact test
** Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Correlations

The sponsor reported correlations of MRI parameters (changes in lesion volume; lesion activity) with clinical endpoints (disability progression and exacerbation rate). Generally, the p-values were highly significant; however, the R values were quite small, suggesting that the correlations were not clinically important. R values for correlations between MRI parameters and clinical outcome were generally similar for both treatment groups, and consistently less than 0.30. The largest R value was for a correlation between two MRI parameters (r=0.36 for % change in lesion volume and number of new or enlarging lesions). Salient correlations for the study overall (Betaseron plus placebo subjects) are shown in Table 21.

Depression/MADRS analysis

The MADRS was administered quarterly by the Treating Physician. Higher scores indicate greater depression, with the highest scores indicative of suicidal risk. There was minimal change in the scores over time in both groups. Proportions of subjects with moderate or severe depression were distributed fairly evenly across groups (n=22 for placebo, n=17 for interferon). Seven (7) subjects in the placebo group and 2 subjects in the interferon group had scores indicating suicidal risk at some time during the study.

Table 21: Sponsor's correlations between MRI parameters and clinical outcome – (placebo and Betaseron subjects combined)

	r	95% CI*	p-value**
Percentage change in lesion volume at endpoint *and*:			
♁ – EDSS at endpoint	0.17	[0.11; 0.23]	0.0000
Time to confirmed progression	0.11	[0.05; 0.17]	0.0005
Annual relapse rate	0.14	[0.08; 0.20]	0.0000
Number of new or enlarging lesions	0.36	[0.31; 0.41]	0.0000
Percentage change in lesion volume during Year 1 *and*:			
– EDSS at endpoint	0.11	[0.05; 0.18]	0.0003
Annual relapse rate	0.09	[0.03; 0.15]	0.0019
Active lesions Months 1-6 and 19-24 *and*:			
♁ Annual relapse rate through Year 2	0.21	[0.05; 0.36]	0.0080
– EDSS at Year 2	-0.02	[-0.19; 0.15]	0.8420
% Change in Lesion Volume at Year 2	0.30	[0.18; 0.43]	0.0000
Number of scans with new active lesions *and*:			
Annual relapse rate through Year 2	0.27	[0.12; 0.42]	0.0005
Active lesions Months 1-6 *and*:			
Annual relapse rate through Year 2	0.27	[0.11; 0.43]	0.0010

* Goodman - Kruskal correlation coefficient; ** Based on overall correlation and SE
♁ Prospectively-defined primary endpoints

Exploratory analyses on depression/MADRS

There has been some concern regarding transient depression upon initiation of Betaseron treatment. Thus, for the complete dataset, CBER analyzed the change in total MADRS score between baseline and Month 3 in 3-point categories. No trend suggestive of a differential effect emerged. CBER also analyzed the changes in MADRS scores by treatment within each 3-month interval, and there was no evidence of a significant treatment difference.

Neutralizing antibodies

Serum was assayed for NAB at baseline, months 1, 2, 3 and every 3 months thereafter. Subjects were considered to be NAB+ if they had quantifiable titers at two consecutive study visits. In addition, the sponsor performed exploratory analyses using titer cutoffs of 1:20, 1:100 and 1:400. Using the above definition for identifying NAB+ subjects and a titer cutoff of 1:20, the false positive rate was very low, with only three (3) subjects in the placebo group (1%) classified as positive. In the interferon group, there were 100 subjects who developed NAB+ status at some time during the study (28%). Seventy percent (70%) of the interferon subjects who eventually became NAB+ did so by their Month 6 visit. The sponsor noted that the persistence of NAB positivity was low, in that 50 (50%) of NAB+ sponsor subsequently had non-quantifiable titers.

The sponsor performed multiple analyses to assess the association of NAB status and treatment response by comparing interferon subjects who were eventually NAB+ to those who were never NAB+. Outcome measures assessed included change in EDSS from baseline, annualized relapse rate, percentage change in T2 volume from baseline, and (in the frequent MRI sub-group) number of newly active lesions and proportion of active scans. There were no statistically significant differences, with the exception of the MRI parameters. The T2 lesion volume changes were significantly higher for the eventually NAB+ sub-group. Conversely, the eventually NAB+ subjects appeared to have less disease activity on the basis of the number of newly active lesions and proportion of active scans.

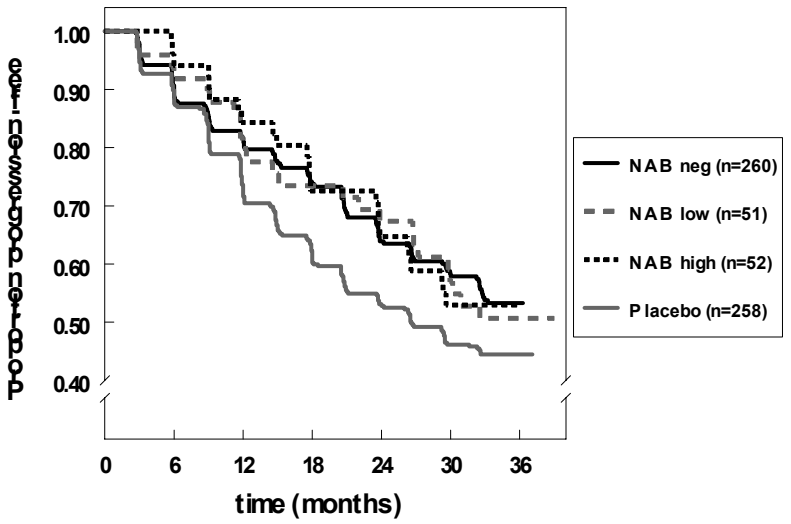
The sponsor performed additional analyses to directly address the question of whether the *change* to NAB+ status is associated with a decrease in efficacy, based on longitudinal data of the interferon subjects who switched from NAB- to NAB+ status during the course of the trial. This allowed each subject to serve as his/her own control for the purpose of assessing whether the change in status was associated with a change in response. The analyses were based on the generalized estimating equations approach to longitudinal data analysis.

The change in EDSS score from baseline was analyzed incorporating a linear time trend. There was no evidence of significant effect of becoming NAB+ with respect to EDSS. With respect to relapses and MRI parameters, however, the results were mixed. A change from NAB- to NAB+ was associated with a 45% increase in relapse rate (p=0.009). The analysis of change in T2 volume suggested a beneficial effect of developing NAB positivity. Conversely, there was a trend suggesting a detrimental effect of becoming NAB+ on the probability of having an active scan (p=0.06).

Exploratory analyses on effect of neutralizing antibodies

Using the complete dataset, CBER performed an exploratory time to event analysis for subjects in the active treatment arm divided into three groups based on NAB status: 1) subjects never NAB+ (NAB neg, n=260); subjects with low NAB titers (n=51); and subjects with high NAB titers (n=52). The results of this analysis are shown in Figure 18. Consistent with the conclusion of the sponsor (based on the interim data), there is no apparent effect of NAB status on the primary efficacy endpoint.

Figure 18: NAB status and progression



CBER also assessed the effect of NAB positivity with regard to the secondary efficacy endpoints. There was no apparent relation between NAB status and annual exacerbation rate (Table 22). There was a trend towards worsened MRI outcomes with higher NAB titers (Table 22); however, the reverse trend was

Table 22: NAB status and secondary endpoints

	Placebo	Interferon		
		Never NAB +	NAB (low titer)	NAB (high titer)
total subjects (N)	358	260	49	51
exacerbation rate ($year^{-1}$) (mean ± SEM)	0.70 ± 0.06	0.51 ± 0.05	0.44 ± 0.10	0.47 ± 0.08
Newly active MRI lesions through Month 36 (mean ± SEM)	8.14 ± 0.52	2.99 ± 0.45	3.59 ± 0.70	7.56 ± 1.39
(N)	345	251	49	50
T2 volume at Month 36, % change from baseline (mean ± SEM)	16.0 ± 2.0%	-4.3 ± 1.8%	-2.0 ± 2.8%	10.9 ± 4.1%
(N)	274	206	40	47

observed with respect to the secondary time-to-wheelchair-bound endpoint. Comparing all NAB-positive interferon-treated subjects (both low and high titer) with NAB-negative interferon-treated subjects, there was a trend towards *delayed* time to wheelchair-bound in NAB-positive subjects (data not shown; p=0.10, log-rank).

In light of the negative results of this analysis with respect to the primary efficacy endpoint, and given the inconsistent results with respect to the secondary endpoints, CBER finds no compelling evidence of an association (positive or negative) between NAB status and efficacy.

Exploratory Analyses for Differential Efficacy

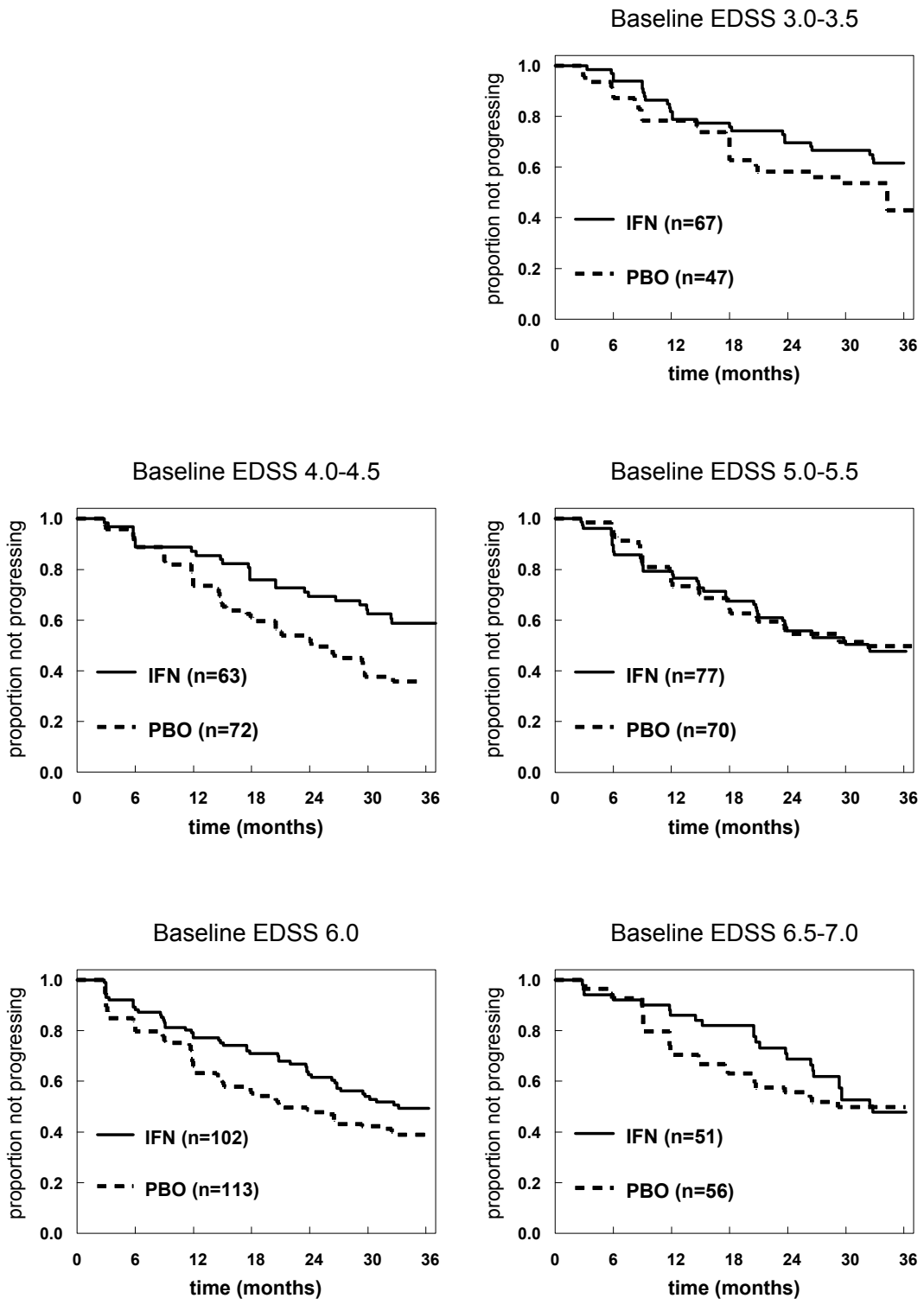
Note: All analyses in this section were performed by CBER. For the purpose of these analyses, one year equals 365.25 days; one month equals one twelfth of a year.

Exploratory analyses were conducted to examine the data for indications of patient subsets that may not have derived benefit from interferon administration, while still incurring adverse effects. Such patients would have an unfavorable risk to benefit comparison. Outcomes assessed were limited to disability progression, time to wheelchair-bound, exacerbation rates and changes in T2 volume.

1. Subgroups defined by baseline EDSS score

A key question is whether the salutary effect of interferon on progression is generalizable to subjects at all levels of baseline disability. Subjects were divided by baseline EDSS scores in one-point increments, and Kaplan-Meier curves were constructed for each sub-group (Figure 19). The subgroup with baseline EDSS = 6.0 was analyzed alone, because of its larger *N*. With the exception of the baseline EDSS 5.0 to 5.5 category, a treatment effect is apparent within all baseline EDSS categories. The apparent lack of effect in the EDSS 5.0-5.5 subgroup would be of concern if there were a trend(s) towards diminishing interferon effects at adjacent higher or lower baseline EDSS; however, these data show strong treatment effects within baseline EDSS categories bracketing the 5.0 to 5.5 group. Thus, this lack of effect in a single category is more likely a chance occurrence and is not a matter of important concern. The effect of interferon on progression appears to be generalizable across all baseline EDSS categories from 3.0 to 6.5.

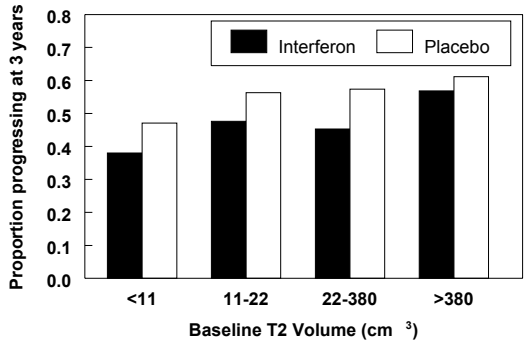
Figure 19: Time to Progression by Treatment Arm and Baseline EDSS Status



2. Subgroups defined by baseline T2 volume

T2 volume is thought by some to provide an assessment of overall disease burden. Thus, analyses of efficacy outcomes across T2 volume provide another measure of generalizability across the spectrum of MS severity. CBER assessed time to progression in interferon-treated subjects, with subgroups divided by baseline T2 volume in quartiles (<11 cm^3, 11-22 cm^3, 22-380 cm^3 and >380 cm^3, Figure 20). Within both treatment groups there is generally a positive relation between baseline T2 volume and overall progression at three years. More importantly, there appears to be an interferon treatment effect across all quartiles of baseline T2 volume, supporting generalizability of the treatment effect to all baseline levels of MRI-defined disease.

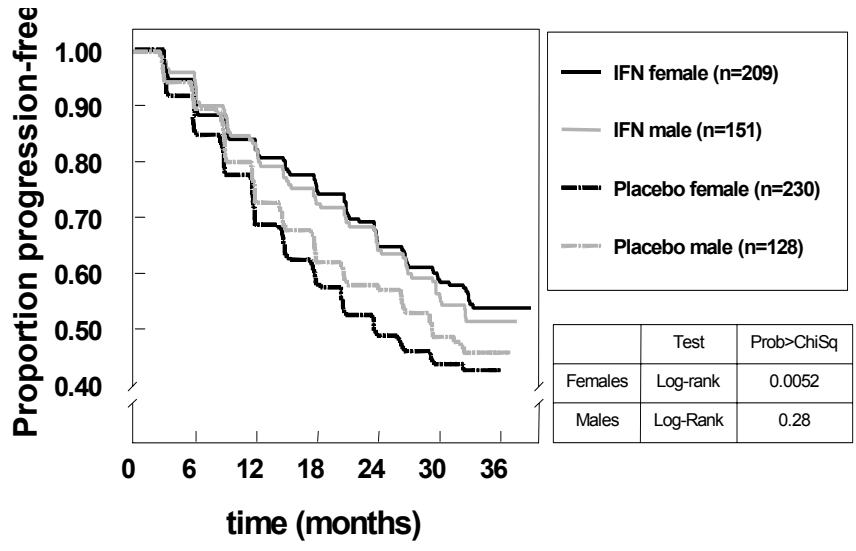
Figure 20: Progression by treatment and baseline T2 volume



3. Subgroups defined by gender

The predilection of MS for women of childbearing age suggests that some aspect of its pathobiology is gender- or hormonally-related. CBER performed analyses of time to progression and time to wheelchair-bound with subgroups defined by gender. Whereas the treatment effect for females was robust, the effect in males was not statistically significant (P log rank = 0.0052 for females, 0.28 for males, Figure 21), though it remained directionally in favor of interferon. When simple proportions were analyzed using Fisher's Exact Test, the overall progression rates in females were 44 and 55% in the interferon and placebo arms, respectively (p=0.0218). For males, however, the rates of progression were 47% and 52%, respectively (P=0.47).

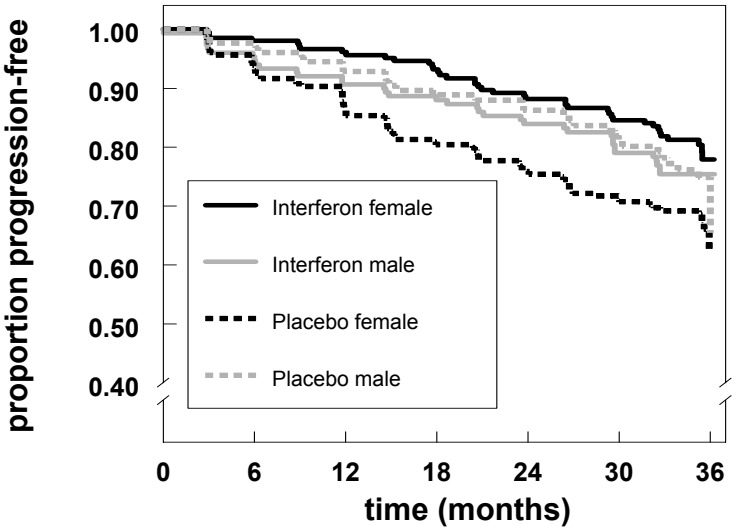
Figure 21: Time to progression by gender



	Test	Prob>ChiSq
Females	Log-rank	0.0052
Males	Log-Rank	0.28

For the secondary time to wheelchair-bound endpoint, a gender difference was even more apparent (Figure 22). Whereas there was a strong treatment effect in females, in males there was a slight trend in favor of *earlier* confinement to wheelchair in interferon-treated subjects. Overall, both groups of males were generally as likely to avoid wheelchair confinement as the female subgroup of the active treatment arm. This could be a chance finding, or a manifestation of greater upper body strength in males (versus females), which could serve to delay the need for wheelchair confinement.

Figure 22: Time to wheelchair-bound by gender



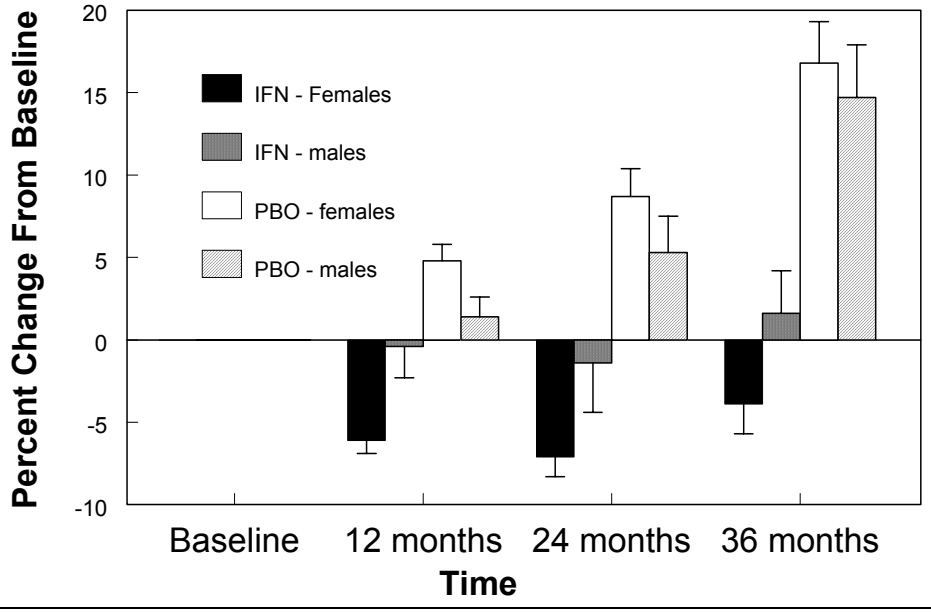
CBER also analyzed annualized relapse rates by gender (Table 23). There was a highly significant interferon treatment effect in females; however, the effect was not statistically significant in males (although it was directionally similar to that of females). Within the placebo arm, males tended to have a lower relapse rate than females.

Table 23: Annualized relapse rate by gender

	Males		Females	
	Placebo	Interferon	Placebo	Interferon
N	128	151	230	209
mean	0.62	0.51	0.74	0.48
SEM	0.09	0.07	0.07	0.05
p-value*	0.22		0.008	

The MRI T2 volume data provide the most convincing evidence of a salutary interferon treatment effect in males. On this outcome measure, there was a striking treatment effect for both males and females (Figure 23).

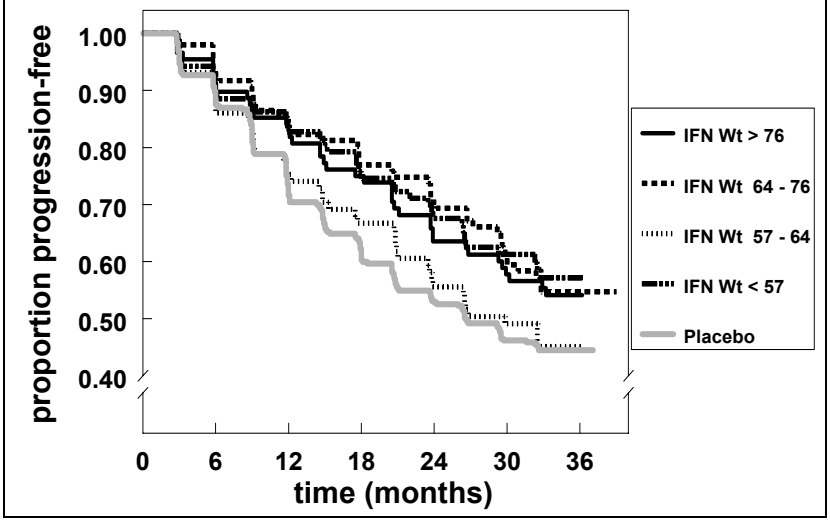
Figure 23: Change in T2 lesion volume by year - analysis by gender



4. Subgroups defined by weight

Betaseron was not administered on a weight-adjusted basis; therefore, CBER considered the possibility of a relation between subject weight and efficacy. Such a relation, if borne out by the data, could explain differential efficacy in males and females. CBER performed a time to event analysis to explore a possible relation between subject weight and efficacy. Subjects in the active treatment arm were grouped by body weight quartiles and compared to controls (Figure 24). There appeared to be similar efficacy in 3 of 4 quartiles, with no effect in the second lowest weight quartile (57-64 kg). The lack of an apparent relation between weight and efficacy suggests that the "one dose fits all" dosing approach is reasonable. Moreover, it suggests that fixed dosing is not responsible for a decrease in efficacy in males.

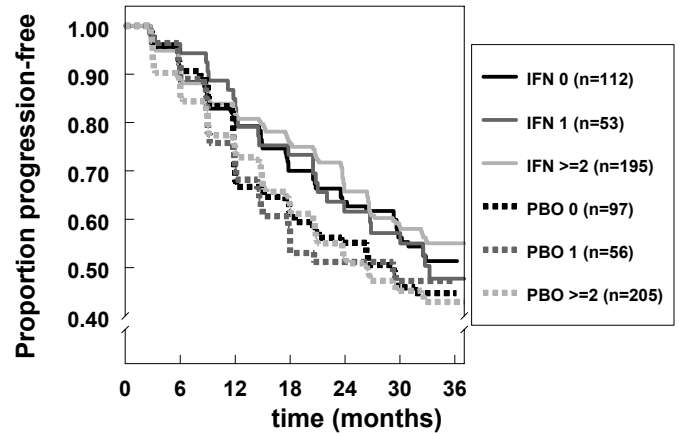
Figure 24: Time to progression with subgroups defined by baseline weight



5. Subgroups defined by number of exacerbations in two years prior to enrollment

Exploratory analysis #4 (page 39) was intended to address the question of whether interferon delays disease progression in a "pure" secondary progressive sub-population of MS, and to determine whether the salutary effect of interferon on disease progression occurs independently of its effect on relapses. CBER performed a variation on that analysis, to assess time to progression in subgroups divided by number of exacerbations in 24 months prior to enrollment (none, 1, ≥2). As shown in Figure 25, the effect of interferon on time to progression appears to be independent of the number of relapses in the 24 month pre-study period. In addition, there appears to be no important interaction between pre-study exacerbations and time to progression in the placebo arm, suggesting that exacerbation rate is not predictive of future progression in MS.

Figure 25: Time to progression with subgroups defined by relapses in 2 years prior to study



CBER also analyzed annualized relapse rate on study within these subgroups (0, 1, or ≥2 relapses in the 24 month pre-study period, Table 24). The number of relapses in the pre-study period appears to be

predictive of subsequent relapses in both interferon and placebo groups. Within each sub-group, a significant treatment effect is apparent.

Table 24: Annual relapse rate by pre-study relapse rate

	Exacerbations in 24 months before study	Treatment group: Placebo	Treatment group: Interferon
N			
	0	n = 97	n = 112
	1	n = 56	n = 53
	≥ 2	n = 205	n = 195
Annualized exacerbation rate			
	0	0.32 ± 0.05	0.19 ± 0.03
	1	0.61 ± 0.17	0.43 ± 0.06
	≥ 2	0.91 ± 0.08	0.69 ± 0.06

Safety

Deaths

There were four deaths during the study; two of which were completed suicides. Summary by treatment group:

Interferon group: A 44 year-old female died of virulent bronchopneumonia at Week 16 on study. A 49 year-old female who had been on interferon for 18 months died of a massive pulmonary embolus, two months after interferon was discontinued. This subject had multiple risk factors for embolism.

There were two attempted and one completed suicide in the active treatment arm. Two of these patients were taking interferon at the time of the attempt (one for two months, the other for 21 months). The third subject had been on interferon for 33 months, but the agent had been discontinued 6 weeks earlier for an adverse event.

Placebo group: There were four attempted suicides and one completed suicide; there were no other deaths.

Serious Adverse Events (SAEs)

For the interferon arm, a total of 503 SAEs were reported in 179 subjects (49.7%). For the placebo arm, 678 SAEs were reported in 192 subjects (53.6%). The most frequent SAE was MS exacerbation requiring hospitalization. Excluding all MS exacerbation-related events, there were 195 and 199 SAEs in the interferon and placebo arms, respectively. SAEs unlikely to be related to MS exacerbations with occurrences in at least 3 subjects, as well as SAEs that occurred in 2 interferon-treated subjects and no placebo-treated subjects are summarized in Table 25.

Injection site problems were significantly associated with interferon use. In the interferon group, 4 SAEs were classified as injection site reactions, of which 2 required discontinuation of study agent. There were 4 instances of injection site necrosis classed as SAEs; 3 required discontinuation and 1 required interruption of study agent. There were also 3 instances of injection site abscesses; 2 required discontinuation of study agent. There were no injection site-related SAEs in the placebo group.

Reviewer's Comment: Upon review of the SAE narratives, the distinctions between injection site reaction, necrosis and abscess were not clear. One subject developed a multicentric, necrotizing vasculitis, consistent with an arthus reaction, requiring surgical incision and drainage and parenteral antibiotics. A second subject developed a large area of cellulitis on the thigh, requiring incision and drainage and intravenous antibiotics. A third subject developed an abscess requiring incision and drainage with oral antibiotics. A fourth subject was treated with oral antibiotics for bilateral arm

abscesses. An additional subject, with an ulceration that failed to completely resolve after treatment with topical antibiotics, subsequently developed a radial sensory neuropathy as a complication, treated with local antibiotics and steroids. Thus, the overall incidence of serious injection site reactions is probably in the 1-2% range.

The data suggest excess cardiovascular morbidity in interferon-treated subjects (due to myocardial infarction and pulmonary embolism), although the total number of subjects with cardiovascular SAEs is limited. No other trends are apparent in the data. The potential for increased incidence of infection is a concern with use of this immunomodulatory agent. Although there appears to be a trend in favor of increased pyelonephritis and cystitis in the active treatment arm, the overall incidence of urinary tract infection is similar in the two groups. Pneumonia, infection, and pharyngitis are fairly equally distributed between treatment arms. The use of systemic antibiotics provides an index of susceptibility to infection, and is reviewed in concomitant medications.

Table 25: Notable serious adverse events (total events/treatment group)

Symptom/Sign	Interferon	Placebo
Death/Suicide		
Death (non-suicide)	2	0
Suicide	1	1
Suicide attempt - incomplete	2	4
Infection		
Urinary tract infection	11	10
Pyelonephritis	3	1
Cystitis	2	0
Pneumonia	4	3
Fever	7	3
Infection	2	6
Flu syndrome	4	5
Pharyngitis	2	0
Injection Site		
Abscess	3	0
Necrosis	4	0
Reaction	4	0
Laboratory Abnormalities	5	3
Cardiovascular		
Myocardial infarct	2	0
Chest pain	1	2
Deep thrombophlebitis	2	2
Pulmonary embolus	3	0
Urinary Tract		
Urinary tract disorder	6	7
Urinary incontinence	0	6
CNS		
Depression	7	10
Vertigo	3	5
Dizziness	1	6
Convulsion	3	3
Emotional lability	0	5
Somnolence	1	2
Deafness	0	3
Gastrointestinal		
Gastrointestinal disorder	1	3
Constipation	3	2
Diarrhea	2	1
Fecal incontinence	1	2
Anorexia	1	3
Nausea	3	3
Vomiting	4	5
Musculoskeletal		
Arthralgia / arthritis	5	0
Back pain	5	5
Pain in extremity	2	3
Miscellaneous		
Surgery	7	13
Accidental injury	9	6
Pain	4	6
Bone fracture (not spontaneous	4	3
Spontaneous bone fracture	1	4
Cataract	2	1
Dyspnea	2	0
Eye pain	2	0

Non-serious Adverse Events (Table 26)

When adverse events significantly associated with Betaseron use in the present study are compared to the Adverse Reactions section of the existing labeling for Betaseron for relapsing remitting MS, these events fall into three categories:

1) significant association with Betaseron use in present RRMS labeling: flu-like syndrome, injection site reaction, injection site necrosis, myalgia and leukopenia
2) included in RRMS labeling, but at a level that was not statistically significant: hypertonia, hypertension and abdominal pain
3) events not included in present labeling: abscess and rash

Because Betaseron use has been associated with injection site necrosis and abscess formation, CBER analyzed injection site necrosis, ulceration and abscess formation by location. Abscess was reported as an event in 22 subjects. In 7 subjects, abscesses were reported at sites that suggested a relation to study agent injection: 6 were reported in the interferon arm and 1 in the placebo arm. In 15 subjects (10 interferon subjects, 5 controls), abscesses were located at sites deemed likely to be unrelated to test agent injection (dental, perianal and axillary abscesses). Injection site necrosis/ulceration was reported in 19 interferon-treated subjects and no control subjects. Severity was reported as mild in 6 subjects, moderate in 10 subjects and severe in 3 subjects. Injection site necrosis necessitated treatment discontinuation in 1 subject and treatment interruption in 3 subjects.

Laboratory Abnormalities

Elevations in hepatic transaminases and hematologic abnormalities occurred more frequently in interferon-treated subjects. Grade 3 and 4 laboratory toxicities necessitated study drug dose interruption (except grade 3 lymphopenia) and totaled 51 and 119 in the placebo and interferon groups, respectively. Twenty-three percent (23%) of interferon subjects experienced grade 3 or 4 lymphopenia (absolute lymphocyte count $\leq 0.74 \times 10^9$/L), whereas only 8% of placebo subjects experienced grade 3 or 4 lymphopenia. Grade 3 and 4 toxicities requiring interruption of study medication occurred in 27 placebo subjects (7.5 %) and 52 interferon subjects (14.4%).

Other Abnormalities

Rash was reported in 73 subjects in the interferon group (20%) and 43 subjects in the placebo group (12%). This difference was statistically significant, and is notable because rash was not significantly associated with interferon use in the RRMS trial.

Seizures occur with increased incidence in MS patients. In this study, seizures were fairly evenly distributed between treatment arms, occurring in 8 subjects in the interferon arm and 6 in the placebo arm. Seizure events were classed as serious in 3 subjects in each arm.

Table 26: Non-serious Adverse Events (number of subjects with event)

Symptom/Sign	Interferon		Placebo	
	number	(%)	number	(%)
Significantly Associated with Interferon				
Flu syndrome	220	(61.1)	141	(39.4)
Hypertonia	147	(40.8)	112	(31.3)
Injection site reaction	165	(45.8)	35	(9.8)
Injection site inflammation	173	(48.1)	14	(3.9)
Myalgia	84	(23.3)	33	(9.2)
Rash	73	(20.3)	43	(12)
Abdominal pain	40	(11.1)	23	(6.4)
Leukopenia	37	(10.3)	18	(5)
Injection site necrosis/ulceration	19	(5.3)	0	(0)
Injection site abscess	6	(1.7)	1	(0.3)
Abscess, other	10	(2.8)	5	(1.4)
Hypertension	16	(4.4)	6	(1.7)
Chills or fever	231	(64.2)	75	(20.9)
No Significant Association with Interferon				
Asthenia	225	(62.5)	207	(57.8)
Headache	170	(47.2)	145	(40.5)
Myasthenia	140	(38.9)	143	(39.9)
Neuropathy	135	(37.5)	148	(41.3)
Paresthesia	125	(34.7)	141	(39.4)
Abnormal gait	122	(33.9)	121	(33.8)
Rhinitis	98	(27.2)	113	(31.6)
Depression	96	(26.7)	112	(31.3)
Pain	110	(30.6)	89	(24.9)
Back pain	92	(25.6)	86	(24)
Urinary tract infection	78	(21.7)	91	(25.4)
Ataxia	70	(19.4)	81	(22.6)
Arthralgia	72	(20)	72	(20.1)
Pharyngitis	58	(16.1)	69	(19.3)
Accidental injury	50	(13.9)	61	(17)
Dizziness	50	(13.9)	51	(14.2)
Pain in extremity	52	(14.4)	44	(12.3)
Nausea	46	(12.8)	48	(13.4)
Constipation	43	(11.9)	44	(12.3)
Infection	46	(12.8)	39	(10.9)
Incoordination	38	(10.6)	45	(12.6)
Urinary incontinence	30	(8.3)	52	(14.5)
Abnormal vision	34	(9.4)	42	(11.7)
Bronchitis	32	(8.9)	41	(11.5)
Insomnia	44	(12.2)	29	(8.1)
Vertigo	29	(8.1)	42	(11.7)
Emotional lability	28	(7.8)	40	(11.2)
Paralysis	27	(7.5)	35	(9.8)
Urinary tract disorder	26	(7.2)	35	(9.8)
Amblyopia	26	(7.2)	33	(9.2)
Diarrhea	24	(6.7)	35	(9.8)
Cystitis	26	(7.2)	32	(8.9)
Somnolence	28	(7.8)	30	(8.4)
Diplopia	24	(6.7)	31	(8.7)
Cough increased	17	(4.7)	37	(10.3)
Tremor	23	(6.4)	31	(8.7)
Peripheral edema	26	(7.2)	26	(7.3)
Urinary urgency	27	(7.5)	24	(6.7)
Injection site pain	31	(8.6)	19	(5.3)
Malaise	28	(7.8)	19	(5.3)

Depression is not uncommon in MS patients, particularly in a patient population with more advanced disease (as in this study). Any therapy that can exacerbate this propensity has the potential to convey significant harm to these patients. Depression and suicidal tendencies are reported in the labeling as adverse reactions for Betaseron, identified in the Warning section. As noted above, there were three suicide attempts in the interferon arm (one completed), and five suicide attempts in the placebo arm (one completed, p=NS). Depression was reported as a Serious Adverse Event in 7 interferon-treated subjects and 10 subjects in the placebo arm.

Additional analyses were performed on the complete datasets to assess the incidence of depression as an adverse event. When reported depression was analyzed without regard to severity, there was no apparent difference between treatment groups. An additional analysis was performed to assess proportions of subjects with mild, moderate and severe depression. Again, there was no differential effect with respect to interferon use.

Finally, as an indirect assessment of depression, CBER analyzed antidepressant use in the complete concomitant medication dataset. During the first year on study, there appeared to be slightly greater use of antidepressant medications in the placebo group, whereas this trend was reversed during the second year. These differences were not statistically significant. During the final year, antidepressant use was fairly evenly distributed.

Taken together, the results of the CBER analysis are in agreement with those of the sponsor, in that we find no apparent association between interferon use and depression.

Menstrual Disorders

Menstrual disorders, including metrorrhagia, amenorrhea, menopause and intermenstrual bleeding were similarly distributed between treatment arms. Metrorrhagia was reported as a severe adverse event in one subject in each treatment group.

Assessment

Study design

This was a randomized, double-blind, placebo-controlled, multicenter, multicountry study designed to assess the efficacy and safety of interferon β-1b in patients with secondary progressive forms of multiple sclerosis. The dose used was the dose licensed for use in relapsing remitting MS, 8 million International Units (0.25 mg) s.c., 3.5 times each week. The primary endpoint was progression of disability as determined by EDSS scores. There were four secondary endpoints. These were not prospectively designated as to importance, and included time to becoming wheelchair-bound, annual relapse rate, percentage change in MRI T2 lesion volume, and number of newly active MRI lesions, Months 1-6, Months 18-24.

Assessments of disablity, relapse rates and MRI outcomes are regarded within the field to be the most important assessments in MS clinical trials. For the sub-set of patients with SPMS, the disability outcome is of greatest import. The primary efficacy endpoint and the secondary endpoint on time to wheelchair-bound directly address the issue of progression.

The secondary exacerbation endpoint in generally important in MS trials; however, interferon is licensed as an agent to decrease exacerbations, and a salutary effect is to be expected, based on prior studies.

MRI is presently unproven as to its clinical meaningfulness; however, it has been utilized frequently in clinical studies as a surrogate of both MS disease burden and activity. The ability to perform and analyze serial MRI studies in an objective and quantitative manner are important advantages of this technique. Thus, the secondary MRI endpoints are viewed as being important supportive information.

The tertiary endpoints were numerous, and not prospectively designated with respect to relative importance. Many were designed to support the primary and secondary endpoints, others have not been properly validated in this patient population and should be considered exploratory.

Study conduct

The study as conducted enrolled 718 subjects. It was terminated early by the sponsor on the basis of a prospectively planned interim analysis of efficacy and safety. Although there was concern in the medical community regarding the possible ramifications of early study termination, the study had achieved 82% of the planned patient-years at the time the database was locked for the interim analysis. Moreover, at the time the study was terminated, it had achieved 92% of its planned patient-year experience. Thus, despite early termination, the study captured nearly all of its planned patient experience. The data as of the interim cut-off date (November 20, 1997) were analyzed by the sponsor for efficacy, whereas the complete dataset was evaluated for safety.

The study treatment was reasonably well-tolerated. Approximately 14.4% of subjects in the interferon group discontinued study treatment because of adverse events; these occurred at a consistent rate of approximately 5% per year. The vast majority of EDSS data were available for analysis. Not including screening and baseline evaluations, or evaluations following early study withdrawal, a total of 8464 EDSS evaluations were planned, of which only 68 (0.8%) were missing.

Randomization and blinding

Randomization was performed centrally. There were no notable imbalances at any site with respect to the numbers of subjects randomized to treatment group.

The study included measures to maintain the treatment blind by administration of ibuprofen prophylactically to reduce flu-like symptoms associated with interferon use. Compliance with prophylactic ibuprofen was only moderately successful, however, with approximately 59% of subjects in the active treatment group and 38% of control subjects using ibuprofen during the first two study months. Moreover, ibuprofen use declined during the first year of the trial, and a consistent 3:2 imbalance in favor of the use of non-steroidal anti-inflammatory drugs in the interferon group persisted throughout the study. These observations suggest subject unblinding.

Although injection site reactions, flu-like symptoms and the need for continuance of non-steroidal anti-inflammatory drugs could serve to unblind subjects and their Treating Physicians, a specially designated EDSS Physician was used to assess EDSS scores for the primary endpoint. These physicians were to remain as uninformed as possible as to any elements of

patient status that might differentiate treatment effects. Because EDSS criteria are largely independent of patient effort variability effects, the primary endpoint is deemed to be largely reliable in this study, even in the presence of unblinding effects of treatment. This component of the study design argues for accepting the assessments as unbiased. The results of analyses of blinding questionnaires tend to support these conclusions.

The adequacy of blinding was assessed by use of three-choice blinding questionnaires ("placebo," "Betaseron," "don't know"), completed by Treating Physicians, EDSS Physicians and patients. Given that all individuals associated with the trial would recognize the importance of the integrity of the blind as it relates to the interpretability of the results, there would be knowledge that a "don't know" response was the preferred choice. Such bias may have produced increased selection of that response. The results of the blinding questionnaire, as reported by the sponsor, are uninterpretable. IN CBER's analysis of the blinding questionnaire results, the elimination of "don't know" responses estimates that Treating Physicians, subjects, and EDSS Physicians guessed correctly 82%, 78%, and 59% of the time, respectively. This suggests substantial unblinding of subjects and Treating Physicians (who are privy both to patient symptoms and potential unblinding effects of laboratory data), with maintenance of the blind for EDSS Physicians. As noted above, blinding of EDSS Physicians is critical for the interpretability of the study, and because EDSS Physicians were not able to accurately guess treatment assignment, these results suggest that the EDSS evaluations for the primary efficacy endpoint can be accepted as unbiased.

Study population

The two treatment groups were reasonably well-balanced for demographics and baseline disease status. There was a greater proportion of females in the placebo group. Because MS tends to follow a more benign course in women, excess females in the placebo arm would be expected to bias the results against the active treatment, and is not considered to be an important limitation. There was a tendency for baseline MS disease status to be slightly worse in the placebo group. This would be expected to bias the results in favor of the active treatment, and would tend to counterbalance the gender difference.

This study was designed to evaluate the efficacy of interferon in subjects with secondary progressive forms of MS. As expected, the patient population tended to be older, with a more prolonged duration of MS and higher baseline EDSS than patient populations in previous studies of subjects with RRMS.

Approximately 40% of subjects in this study lacked documentation of a pre-study increase in EDSS, indicative of the gradual progression in disability characteristic of SPMS. For these subjects, the diagnosis of SPMS was based on the judgement of the investigating physician. For the 60% of subjects who did have documented pre-study progression, a Kaplan-Meier time-to-event analysis demonstrated significant efficacy with respect to time to progression. Thus, although the lack of documentation of chronic progressive status in 40% of subjects represents a significant weakness of the study, there is significant efficacy in the remaining 60% of subjects, despite a reduction in statistical power due to decreased N.

Primary efficacy endpoint

The primary efficacy endpoint was time to confirmed disease progression, defined as a 1-point increase in EDSS confirmed at the next scheduled study visit 3 months later (at least 70 days apart). (A 0.5 point increase defined progression if the baseline EDSS was 6.0 or 6.5.) Because of the requirement for confirmation, the Month 33 visit provided the final opportunity for

progression. Ten (10) subjects had less than 5 months on-study, and therefore no time at risk for a confirmable progression.

The primary endpoint showed a statistically significant delay in time to disease progression in the active treatment arm compared to placebo. In all, there were 193 progressions in the placebo group and 163 in the interferon group. The Kaplan-Meier estimates for progression at three years were 55% placebo, 47% interferon.

Because the study encompassed over 1900 patient-years of experience, the statistical power of the study was robust. As such, modest clinical benefit was demonstrable at a level of statistical difference that was quite impressive. To put the results into perspective, interferon prevented the progression of disability in 8% of subjects over three years - - an absolute reduction in the annual rate of approximately 3%.

Exploratory analyses showed a directionally similar if not statistically significant effect of interferon on the time to progression endpoint, irrespective of the method for handling missing data (subjects lost to follow-up), irrespective of the criterion used for EDSS progression, and irrespective of the exclusion of EDSS determinations during relapses.

The results were generally consistent across sites, and within the countries contributing importantly to overall subject number, with Germany and Italy as exceptions. Efficacy was also demonstrable in the sub-group of subjects who would meet strict diagnostic criteria for SPMS. Moreover, the salutary effect of interferon on time to progression was independent of its effects on exacerbations. Concomitant medication used for symptomatic treatment of MS, although imbalanced in some categories, does not appear likely to have altered the EDSS in a manner that would favor the active treatment arm.

In general, the efficacy of interferon was generalizable across all sub-groups, although the treatment effect was less pronounced in males than females.

There were two concerning findings raised by CBER's exploratory analyses. The first issue relates to time of enrollment. When subjects were artificially subdivided by time of enrollment (first versus second half), there was minimal treatment effect apparent in subjects who enrolled in the second half of the study. This finding remains unexplained at the present time, but could be due to an as yet unidentified difference in baseline demographics, baseline disease activity, MS management or assessment, or simply due to play of chance. The second issue relates to the duration of interferon's treatment effect. There was excess progression of disability in placebo subjects during the first year on-study; however, event rates were essentially identical in the two groups after Year 1. Together, these analyses indicate that the positive efficacy of the study was driven by subjects who enrolled in the first half of the study; moreover, only the data obtained during the first year on study were supportive of an interferon treatment effect. Importantly, although benefit appeared to be limited to 1 year, adverse events, sufficiently severe to warrant discontinuation of treatment, occurred at a constant annual rate of approximately 5%. Therefore, the relation between risk and benefit appears to become less favorable with increasing time.

Secondary efficacy measures

Amongst the many secondary endpoints, time to wheelchair-bound relates to disease progression and is probably most relevant. On this endpoint, there was a significant effect of

interferon. As for the primary endpoint, however, the effect was more apparent in females. For this endpoint in males, there was no demonstrable effect whatsoever.

The other secondary endpoints, exacerbations and MRI assessment of disease burden and activity, currently receive the most attention in the field. By all measures, the effect of interferon on exacerbations was highly significant. As was observed on the progression endpoints, however, the interferon-associated improvement in annualized relapse rate in males did not reach statistical significance (although it was directionally in favor of interferon).
There was a strong and clear benefit of interferon with respect to the MRI data, and these data provided the strongest indication of a treatment effect in males. Overall, the secondary endpoints were strongly supportive of the demonstration of interferon efficacy in the treatment of MS.

Tertiary efficacy measures

Numerous tertiary endpoints were supportive of the primary and secondary endpoints on disease progression, exacerbations and MRI variables. Other more exploratory measures of effect were less convincing. In particular, there was no clinically meaningful effect of interferon with respect to cognitive function or quality of life. The Global Evaluation of MS showed a statistically significant treatment effect of interferon; however, this very subjective measure could be confounded by unblinding. The analyses on the ambulation index were mixed, and therefore difficult to interpret.

Safety

Interferon was well-tolerated by the majority of subjects. The majority of serious adverse events were hospitalizations for treatment of MS-related problems. Excluding all MS exacerbation-related events, serious adverse events were balanced in frequency between the interferon and placebo arms. The only serious adverse event identified more frequently in the interferon arm was injection site reaction. Approximately 14% of interferon subjects discontinued study prematurely because of adverse events, compared with 8% of placebo subjects.

In general, adverse events were consistent with the known and accepted adverse reactions in RRMS. The exceptions were hypertonia, rash, abdominal pain, abscess and hypertension, which were significantly associated with interferon in this study. Conversely, this study found no increased incidence of menstrual abnormalities associated with interferon.

Depression, particularly transient depression soon after institution of treatment, has been a concern with respect to interferon. Depression was not assessed in the period soon after initiation of interferon treatment in this study, and could not have been captured. With respect to depression in the longer term, the Montgomery and Asberg Depression Rating Scale has not been validated in this patient population, and the results (showing no difference between groups) are difficult to interpret. The more objective endpoints pertinent to depression do not tend to show increased depression associated with interferon administration. Specifically, in the placebo arm, there was a greater number of suicide attempts (5 versus 3), and a greater number of subjects who experienced depression as a serious adverse event (10 versus 7). Thus, although these results can neither refute nor support an excess of transient interferon-induced depression, they do not support a differential effect of interferon on depression in the longer term.

Summary and conclusions

- This was a randomized, double-blind, placebo-controlled multicenter investigation designed to evaluate the efficacy and safety of Interferon β-1b 8 mIU SC on a QOD (alternate day) schedule in patients with chronic progressive forms of multiple sclerosis.

- The primary endpoint was progression of disability, as measured by a 1.0 point change in the Kurtzke Expanded Disability Status Scale (EDSS) score, sustained for at least 6 months. (A 0.5 point increase defined progression for subjects with baseline EDSS $\geq$ 6.0.) This definition was selected as an indication of a clinically meaningful increase in disability, and to minimize variability due to transient exacerbations and/or intra-rater variability.

- Numerous secondary endpoints included time to becoming wheelchair-bound, annual relapse rate, and magnetic resonance imaging (MRI) assessments of disease status (change in lesion volume and number of new lesions).

- The study was concluded early after a planned interim efficacy analysis showed statistically significant results. The completed trial enrolled 718 patients, with both treatment durations and evaluations ranging from 0 to 39 months. The mean time at-risk of progression was 31.3 months.

- The study population consisted of patients with secondary progressive forms of MS with moderate to severe disability (EDSS of 2.0 to 7.0, median 5.5).

- Demographic characteristics were generally balanced between the treatment groups, although there was a trend towards increased females in the placebo group, which would tend to bias the results against the active treatment. Conversely, there were minor imbalances in baseline disease status with respect to EDSS, relapse rates and T2 lesion volume. Although individually small, each of the imbalances in baseline disease status was directionally in favor of more severe disease in the placebo group. Thus, although there were slight imbalances in baseline demographic characteristics and disease status with the potential to bias the study, these imbalances were in opposite directions, such that the study, as a whole, appears to be reasonably well-balanced.

- The primary endpoint of time to progression in EDSS was delayed by treatment with interferon (p=0.0037), with estimated annualized rates of progression of 15.7% and 18.6% in the placebo and interferon groups, respectively. These results were consistent at 5 of the 6 largest centers, with only the Rome site failing to show a trend toward efficacy. There was one notable peculiarity in the primary efficacy data set, in that a differential treatment effect was detected between the first half of the subjects enrolled (in whom there was a robust treatment effect) compared to the second half of subjects enrolled (in whom there was no demonstrable treatment effect). The explanation for this disparity is unknown.

- The interferon treatment effect was largely limited to the first year on-study, whereas adverse events, sufficient to warrant treatment discontinuation, occurred at a constant annual rate of approximately 5%. Thus, the relation between risk and benefit appears to become less favorable with increasing time.

- For the secondary endpoint time to wheelchair-bound, there was a significant effect of interferon with annualized rates of wheelchair-bound of 8% and 11% in the interferon and control groups, respectively (p=0.0053).

- There was a statistically significant effect of interferon with respect to annual exacerbation rate; this was consistent across all levels of relapse severity. Based on an analysis of relapses and time at-risk for individual subjects, annual relapse rates were 0.49 and 0.70 exacerbations per year in the interferon and control groups, respectively.

- There was a highly significant interferon treatment effect with respect to MRI percentage change in T2 lesion volume from baseline to last scan available, with data were collected through 24 months in approximately 85% of subjects.

- An apparent gender-related disparity in treatment effect was observed consistently across most endpoints, with a smaller treatment effect in male subjects. For males, there was a strong trend towards benefit on the primary time-to-event disability endpoint; however, only a marginal treatment effect was observed on the secondary time to wheelchair-bound and relapse endpoints. The data were strongly supportive of a treatment effect in males only on the secondary MRI endpoint (T2 volume). Exploratory analyses of efficacy versus subject weight suggest that the gender-related disparity is not related to disparate weights of male and female subjects.

- Interferon use was associated with a significant reduction in MRI gadolinium-enhancing newly active lesions. The treatment effect was highly significant, both during months 1-6 and during months 19-24. The persistence of biologic activity at the 19-24 month time interval supports durability of the treatment effect.

- Treatments were well-tolerated by most subjects. Fourteen percent (14%) of interferon-treated subjects and 8% of control subjects discontinued the study agent prematurely due to an adverse event. There were three deaths in the interferon group: one death from bronchopneumonia, one from a massive pulmonary embolus and one from suicide. There was one completed suicide in the control group and no other deaths. Most serious adverse events were MS-related hospitalizations. As expected, there was an excess of injection site problems in the interferon group, with 13 serious injection site-related adverse events in the interferon group and none in the control group. Although the overall numbers were small, there appeared to be excess cardiovascular morbidity in the interferon group: myocardial infarction was reported in 2 subjects, and pulmonary embolism was reported in 3 subjects (1 fatal), whereas there were no reports of MI or PE in the control group. In addition, there were 2 subjects with deep thrombophlebitis in the interferon group, versus 1 subject in the control group.

- There was no evidence of increased depression in the interferon group, based on adverse event reports, reported attempted and completed suicide, concomitant anti-depressant use, or the Montgomery and Asberg Depression Rating Scale.

- Side effects typically associated with Betaseron use were observed frequently in the study.

Conclusions:

Based on the results of the multicenter European study of Betaseron in secondary progressive multiple sclerosis, Betaseron appears to be effective in delaying progression of disability in this patient population, with a favorable balance of risk and benefit. The clinical significance of the treatment effect is modest, whereas the level of statistical significance is robust. Presumably, this disparity resulted from a lower than expected dropout rate, which had the effect of statistically overpowering the study. The study succeeded on the secondary endpoints as well, with clear demonstration of a Betaseron-associated decrease in exacerbation rate, and striking reductions in the accumulation of MRI T2 lesion volume and the rate of new MRI lesion formation.

The investigation should be considered in the context of what is presently known and accepted regarding interferons in MS. There is evidence of a salutary effect of two interferons in relapsing remitting multiple sclerosis with two licensed products for this indication (Betaseron and interferon β-1a). It is likely that progression of disability in secondary progressive MS and transient exacerbations in relapsing remitting MS are, at least, related, and, at most, different clinical manifestations of the same fundamental pathophysiologic process. There exists, therefore, a reasonable expectation of efficacy for interferons in preventing disability progression in secondary progressive MS. Although post hoc exploratory analyses on the primary endpoint raise some concerns (gender-related disparity in efficacy and a lack of efficacy in subjects who enrolled in the second half of the study), consistent treatment effects across multiple endpoints provide strong evidence of modest efficacy when the data are considered in their entirety.